The Impact of Assimilating Along-Track SLA Data on Simulated Eddy Characteristics in the Agulhas System

by Marc de Vos

A minor thesis submitted in partial fulfilment of the degree of MSc in Ocean & Climate Dynamics

Supervisor: Dr. Björn Backeberg (UCT)

Co-Supervisor: Dr. François Counillon (NERSC)

UNIVERSITY OF CAPE TOWN

Faculty of Science, Department of Oceanography

February 2016

UCT
MARINE RESEARCH
INSTITUTE

Nansen-Tutu Centre
for Marine Environmental Research

I know the meaning of plagiarism and declare that all of the work in this minor thesis, save for that which is properly acknowledged, is my own.

Signed by candidate

Signed Signature Removed

Marc de Vos

February 2016

Abstract

The Agulhas Current System is a vital element of the global ocean-climate system by virtue of its role in the transfer of energy, nutrients and organic material. In the context of working towards better climate change projections, it is necessary to develop a robust understanding of the complex dynamical mechanisms which facilitate this transfer. Mesoscale cyclonic and anticyclonic eddies transport heat, salt, organic matter and nutrients from the Indian Ocean into the South Atlantic Ocean. In so doing, they are key drivers of the Atlantic Meridional Overturning Circulation (AMOC). As such, it is important that they are adequately simulated by numerical models in order to advance the accuracy of climate prediction. In the absence of spatially and temporally coherent observing systems, numerical models provide the capacity to describe the oceanographic conditions of the region. Given the complexity of the regional dynamics, and the challenges it presents to free-running numerical models, data assimilation is a valuable tool in improving simulation quality. An important step in this continuing process is the objective, quantitative evaluation of model configurations, such that they can be continuously refined.

In this study, the impact of assimilating along-track sea level anomaly (SLA) data is investigated with regard to the simulation of mesoscale eddies in the Agulhas System. Two configurations of a Hybrid Coordinate Ocean Model (HYCOM) configuration are analysed; one free run (hereafter 'Free') and one with along-track SLA data from satellite altimetry assimilated (hereafter 'Assim.') via an Ensemble Optimal Interpolation (EnOI) data assimilation scheme. The results of these two configurations are compared with each other, and against a set of corresponding observational data from satellite altimetry (hereafter 'Aviso'). To this end, an automatic eddy detection and tracking algorithm is implemented, in order to quantify eddy characteristics in a coherent and consistent manner.

Assim. shows global improvements in the simulation of eddy density distribution and dynamics over Free; i.e. better agreement with Aviso. South of South Africa and south of Madagascar, Assim. simulates a more realistic number of eddies than Free, although the number south of Madagascar and in the south-west Indian Ocean remains too low. This is of particular interest, as previous analyses of the mesoscale variability in this region have shown sufficient and indeed excessive eddy kinetic energy (EKE). However, when computed only where eddies are definitively known to exist (as opposed to from continuous fields of sea surface height (SSH)), levels of EKE fall short. This suggests that whilst model configurations are simulating adequate levels of variability, they are not simulating and sustaining eddies realistically in this region.

Root Mean Square Error calculations indicate that the simulation of amplitude, rotational speed, surface eddy kinetic energy and anticyclone vorticity by Assim. shows good improvement. Regions of consistent improvement for these characteristics are those to the southwest and southeast of South Africa, the northerly reaches of the Mozambique Channel, and the coastal region immediately south of Madagascar. Translational (drift) velocities in Assim. show some localised improvements, such as in the western reaches of the Mozambique Channel, but appear less accurate in other areas such as the region south of Madagascar. General translational velocities appear excessive south of South Africa. As a caveat, the translational velocities of 51 %, 43 % and 35 % of Assim., Free and Aviso eddies respectively could not be analysed. This is due to their being inadequately tracked by the automatic tracking scheme, despite having coherent dynamic and geometric properties and is likely due to the temporal resolution of the data.

The simulation of eddy geometry by Assim. is slightly less accurate than by Free. Free shows closer similarity to Aviso in the simulation of eddy radius and area than Assim. In most cases, increased relative error in Assim. is due to the underestimation of radius and area.

This study provides further insight into mesoscale eddy properties themselves, as well as insight into how these properties are resolved by three different products. This is useful in providing recommendations for future configurations. Furthermore, the results of this study indicate that the strategic approach to analysing mesoscale properties is important; analysis of discrete eddy statistics may yield different results to analysis of continuous variable fields normally associated with mesoscale variability.

Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

First and foremost, I would like to extend a sincere vote of thanks to my supervisor, Dr. Björn Backeberg, and my co-supervisor, Dr. François Counillon. Bjorn's vision for the project, supervision of its implementation and valuable advice throughout, have played a key role in the materialisation of this study. For assisting me in navigating the more technical aspects of this research, and his skilful guidance during my exchange in Norway, François's input is greatly appreciated. My time at the Nansen Centre in Bergen has had a significant positive impact on my scientific career. It was an opportunity I remain highly grateful for and something I will look back on fondly. To both, I would like to say a further thank-you for the enthusiasm approachability throughout – factors which made my MSc experience thoroughly enjoyable.

I would also like to thank Dr. Ben Loveday for his work on the eddy detection and tracking algorithm, and Dr. Pierrick Penven for the provision of the original code.

For funding this research, I would like to acknowledge the *National Research Foundation (NRF)* of South Africa for the provision of grant 87698. Thanks also to the *Seasonal to Decadal Changes Affecting Marine Productivity: An Interdisciplinary Investigation (SCAMPI)* project for funding my research visit to *the Nansen Environmental and Remote Sensing Center (NERSC)* in Bergen, Norway. The project operates under the *South Africa-Norway Cooperative Research Programme*. Further thanks to the *Nansen-Tutu Centre for Marine Environmental Research (NTC)* for facilitating my exchange to *NERSC*, and to *NERSC* for hosting me. I appreciate the computer time (project number nn2993k) granted for the running of the HYCOM simulations by the *Norwegian Program for Supercomputing (NOTUR)*.

Finally, I would like to thank my family and friends for their support and good humour throughout the duration of my MSc.

Chapter 1. Introduction

The Agulhas Current is the most energetic western boundary current in the southern hemisphere (Reason, 2001). Moreover, it is one of the most important currents in the global ocean-atmosphere system by virtue of its role in energy transport (Backeberg et al., 2014). The current transports approximately 65 Sverdrup of warm water in a south-westward direction along South Africa's east coast (Reason, 2001; Stramma & Lutjeharms, 1997). South of South Africa, it exhibits an eastward retroflection towards the southern Indian Ocean (Reason, 2001; Lutjeharms & van Ballegooyen, 1988) and develops into the Agulhas Return Current, with distinct hydrographic characteristics to that of the Agulhas Current (Lutjeharms & Ansorge, 2001). An important transfer of warm, salty waters from the south-west Indian Ocean to the south Atlantic Ocean occurs here, and is facilitated primarily by the periodic shedding of Agulhas Rings (upper-mesoscale, anticyclonic eddies) from the Agulhas Retroflection (Lutjeharms & van Ballegooyen, 1988; Richardson, 2007; Backeberg et al., 2014). Further mesoscale perturbations in the region of the southern Mozambique Channel and south of Madagascar have been shown to modulate the frequency of ring-shedding events, which are important drivers of the so-called 'Agulhas Leakage' (Schouten et al., 2002; Penven et al., 2006; Biastoch et al., 2008b; Halo et al., 2014a). These eddies (cyclones or smaller anticyclones) may be up to 400 km in diameter (de Ruijter et al., 2002; Ridderinkhof & de Ruijter, 2003; Schouten et al., 2002). They form primarily in either the upper ocean (0-300 m) or intermediate ocean (800-2000 m), depending on whether they were caused by barotropic or baroclinic instabilities respectively (Halo et al., 2014b). Significant variation exists in the estimation of the ratio of Agulhas waters which flow into the south Atlantic to that which flows back into the south Indian Ocean (Lutjeharms et al., 1992; Lutjeharms & Ansorge, 2001). An enhanced understanding and predictability of the Agulhas Current system would likely prove beneficial to industrial, commercial and recreational interests in the region, as well as in the event of oil spills and harmful algal blooms (Backeberg et al., 2014). Furthermore, Agulhas leakage has itself an important influence on the Atlantic Ocean meridional overturning circulation (AMOC) and hence, on global ocean circulation (Peeters et al., 2004; Biastoch et al., 2008a; Beal et al., 2011). In this way Agulhas leakage has an important bearing on global climate and the variability thereof. An example of one such mechanism exists in recent findings which indicate a warming trend in the Agulhas Retroflection (Rouault et al., 2009), coupled with enhanced mesoscale variability of the greater Agulhas system (Backeberg et al., 2012). This enhanced variability is manifested in increased eddy propagation in the Mozambique Channel and in the Retroflection region (Backeberg et al., 2014), which might in turn cause an intensification of the AMOC.

Further, poleward shifts in the westerly wind belt of the southern hemisphere are hypothesized to allow for increased Agulhas leakage (Biastoch et al., 2009). This is in keeping with the assertion that Agulhas leakage is not necessarily dynamically linked with the main Agulhas Current, showing no response to a strengthening or weakening thereof (Loveday et al., 2014). The mechanisms associated with these hypotheses differ from those which point to a *weakening* of the circulation due to the increased input of fresh water from ice sheet and glacial melting (Biastoch et al., 2008a; Beal et al., 2011). Mesoscale eddies such as Mozambique Rings are also important determinants of the marine ecology in the region; particularly in the western reaches of the Mozambique Channel where seabirds forage at the boundaries of eddies (Weimerskirch et al., 2004). Their impact probably extends further, given their associated high nutrient and low oxygen anomalies (Swart et al., 2010) and their tendency to advect phytoplankton-rich, coastal waters offshore (Omta et al., 2009). Agulhas Rings and smaller mesoscale eddies may also boost primary production by raising nutrients from deep waters to upper ocean levels (Robinson, 1983; Lathuiliere et al., 2010). Accurate simulations of these important elements of the Agulhas Current system depend heavily on the resolution of non-linear dynamics such as inertial mechanisms and ring formation (Beal et al., 2011). However, given the dearth of coherent monitoring systems (Backeberg et al., 2014), as well as lingering inaccuracies in model results due to the complexity of the mesoscale dynamics in question (Biastoch et al., 2008b, Backeberg et al., 2014), concrete conclusions about the magnitude of Agulhas leakage and the variability thereof remain elusive (Backeberg et al., 2014). Furthermore, model simulations of elements of the Agulhas system have been shown to be highly sensitive to changes to some of their technical characteristics. These may include numerical schemes (Backeberg et al., 2009), aspects of the momentum advection scheme (Barnier et al., 2006), smoothness of bottom topography as well as the inclusion of parameterizations such as one for horizontal viscosity (Penven et al., 2006). Thus, given the relevance of a robust understanding of the dynamics associated with the Agulhas system (Reason, 2001; Beal et al., 2011), it is important to continue to simulate regional dynamics, incorporating data assimilation techniques and re-analysis, such that results become increasingly realistic (Backeberg et al., 2014). Data assimilation is a strategic process whereby the most probable model state is computed given a dynamical model and a set of observational data (Counillon & Bertino, 2009a). This approach is beneficial to the forecasting value of model products. Data assimilation is crucial for ocean prediction, as, firstly, many dynamical phenomena are the results of instabilities in the flow, rather than deterministic responses to atmospheric forcing. Secondly, atmospheric forcing contains errors, and thirdly, ocean models are still limited by imperfections such as sub-optimal resolution (Chassignet et al., 2007). Continuous data assimilation constrains models solutions such that forecast accuracy is usefully enhanced (Backeberg et al., 2014).

The aim of this investigation is to provide insight into the results of assimilating satellite altimetry data into a HYCOM simulation. As such, the study forms part of the evaluation process of one approach to modelling an important aspect of the global ocean-climate system. Whilst the results themselves do provide stand-alone value in quantifying elements of Agulhas leakage, their real value lies in fostering a clearer understanding of the effects of data assimilation. The investigation should provide some clear directives for future simulation experiments using the HYCOM/data assimilation approach, by identifying and highlighting drawbacks of the system as implemented here.

This manuscript opens with a review of the relevant literature in Chapter 2, concerning recent assessments of Agulhas leakage and mesoscale eddies in the domain, as well as data assimilation. Chapter 3 presents the methodologies employed in this study, including those concerning the dynamical model, observational data and data assimilation, as well as the eddy detection and tracking algorithm. It also describes how eddy properties were calculated. Chapter 4 presents the results of this study, along with a discussion of those results, whilst Chapter 5 presents conclusions.

Chapter 2. Literature Review

2.1 Recent Assessments of Discrete Eddy Dynamics and Agulhas Leakage

Numerous attempts have been made to quantify Agulhas leakage in the Cape Basin (Loveday et al., 2014). Few studies, however, have focussed on the prevalence of mesoscale eddies as a point of departure for such assessments. Many utilise a Lagrangian approach, deploying virtual floats into the Agulhas Current and assessing how many of these cross a defined leakage section. Such approaches may suffer inaccuracies resulting from the exclusion of diffusive processes (by design, Agulhas leakage is limited to a proportion of the main current which is exported to the neighbouring basin) and reduced precision in turbulent areas (Loveday et al., 2014).

Other approaches utilise Eulerian analyses of defined transects. Imposed water-mass criteria and definitions, and the location and geometry of transects subject these approaches to similar limitations in accuracy. These limitations are particularly substantial in eddy-rich regions, of which the Cape Basin is certainly one (Loveday et al., 2014).

As critical determinants of Agulhas leakage and interbasin exchange, mesoscale dynamics must necessarily be accurately resolved before robust results can be achieved.

The work of Richardson (2007) uses a Lagrangian approach to quantify Agulhas leakage, and includes an assessment of the discrete number of mesoscale features. A range of sub-surface floats and surface drifter data is analysed. This includes 321 drifter trajectories for the period 1994-2004. The drifter data was acquired from the Global Drifting Buoy Data Assembly Center at the NOAA Atlantic Oceanographic and Meteorological Laboratory in Miami, Florida. Subsurface float data were acquired from the WOCE Subsurface Float Data Assembly Center. These include data from RAFOS (ranging and fixing of sound) and ALACE (Autonomous Lagrangian Circulation Explorer) floats. The RAFOS floats acquired an approximate cumulative total of 104 years' worth of data, between March 1997 and September 1999. The ALACE floats collected an approximate cumulative total of 86 years' worth of data, between 1996 and 2000 (Richardson, 2007). An assessment of the number of Agulhas rings (anticyclones) and cyclones is made by means of visual analysis of trapped floats' and drifters' tracks. Looping tracks are considered evidence of mesoscale eddies (Richardson, 2007). This kind of analysis yields useful insight into the properties of those eddies which manage to entrain floats or Lagrangian devices, and may give some indication of the number of eddies present around the time when the devices were released. However, the quantification of the number of mesoscale features remains far from robust, as it depends on devices being entrained at all, and the maximum number of eddies identified is constrained unnaturally by the number of available devices and

their spatial distribution relative to eddies. Furthermore, visual identification of eddies, coupled with manual drawing of their tracks introduces a degree of subjectivity to the analysis. Figure 1 shows the eddies identified and tracked from the relatively extensive temporal and spatial Lagrangian data assemblage utilised.

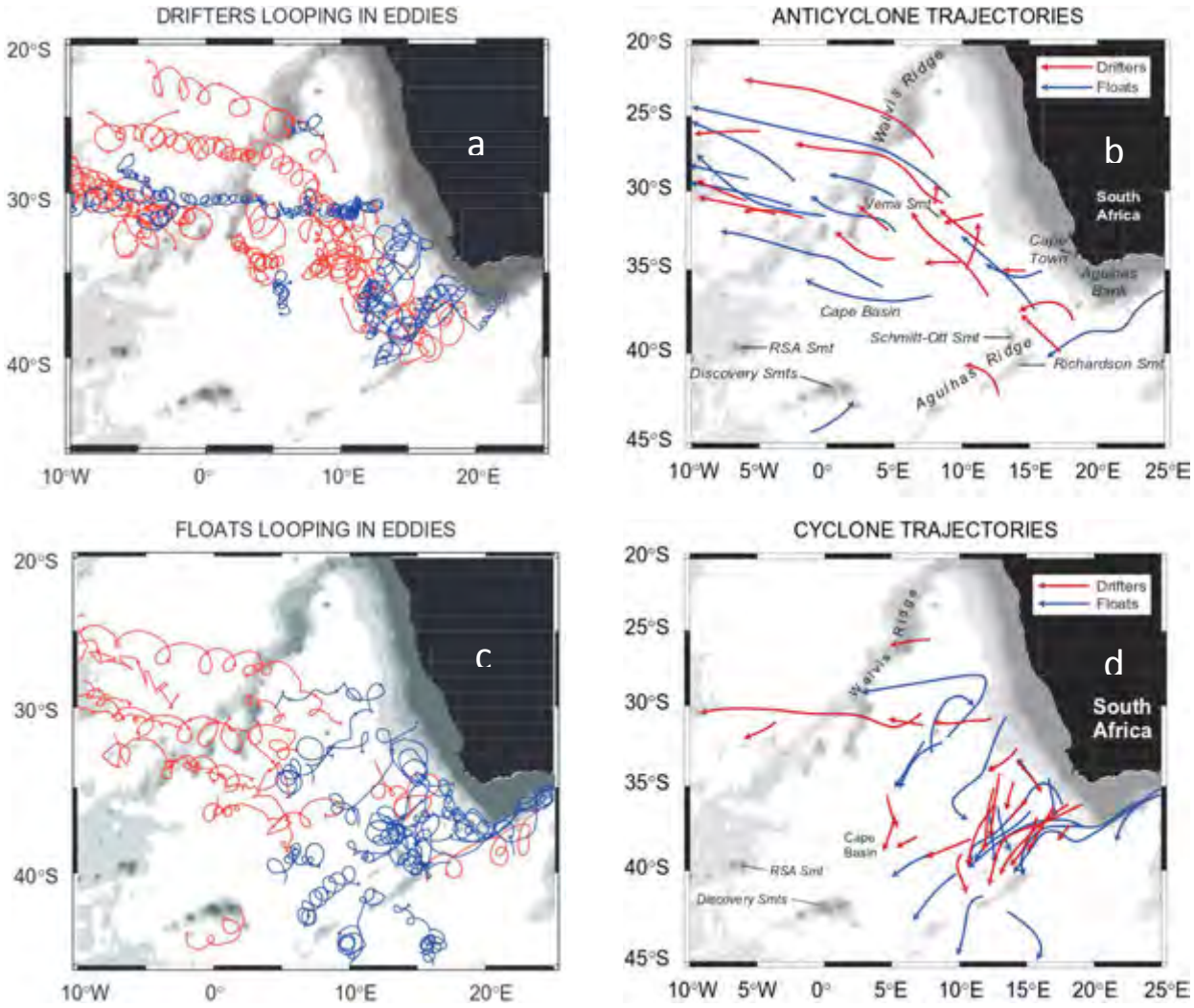

Figure 1. (a, c) Trajectories of surface drifters (a) and floats (c) looping in discrete eddies. Blue trajectories indicate clockwise rotating cyclones and red trajectories indicate counter-clockwise rotating anticyclones. Red trajectories are interpreted to have been in (mainly) warm-core Agulhas rings, and the most southern group of blue trajectories in (mainly) Agulhas cyclones. Some additional, shorter, trajectories in the same eddies were omitted to reduce clutter as were a few short nearly stationary loopers. A few trajectories extended west of the left edge of the figure (10 W). (b, d) Trajectories of anticyclones (b) and cyclones (d) drawn manually through looping drifter and float trajectories. Figure taken from Richardson, 2007.

Numbers of eddies				
	Drifters	RAFOS floats	ALACE floats	Total
Cyclones	17	17	0	33
Anticyclones	24	14	3	35
Total	41	31	3	68

Totals were reduced because one cyclone and six anticyclones were tracked by combinations of drifters and floats.

Table 1. Numbers of cyclones and anticyclones as identified manually from drifter and float trajectories for all data. Table taken from Richardson, 2007.

Nearly all anticyclones exhibit a generally north-westward propagation, with the longest-lived anticyclones (lifetimes > 1 year) having a mean propagation speed of 3.8 km.day^{-1} (Richardson, 2007). The most energetic anticyclones are found in the south-eastern regions of the Cape Basin. Rotational speeds in this region fall into a general range of 40-60 cm.s^{-1}, and generally decrease with distance from the Agulhas retroflection. A mean anticyclone diameter of 200 km and a maximum diameter of 250 km is noted (Richardson, 2007).

Amongst other properties, the work of Backeberg et al. (2014) and Loveday et al. (2014) utilises continuous fields of SSH in order to assess regional eddy kinetic energy (EKE) as resolved by observational and model products. EKE provides a spatial overview of mesoscale variability (Backeberg et al., 2014), much of which is associated with eddies in this region. Of course by computing EKE from continuous data fields (as opposed to only where eddies are present), some EKE might be evident even where an eddy does not exist, if there is some general variance in the mean flow.

Backeberg et al., (2014) make use of two HYCOM configurations; a free run and an assimilated run; as well as drifters from the Global Drifter Program (GDP) and satellite altimetry data from Aviso. The HYCOM configuration with assimilated along-track SLA shows good improvement in the simulation of EKE when compared to observational data from drifters. It is noted that the EKE spatial distribution and pattern computed from Aviso satellite data appears to be more realistic, but underestimates magnitude by a factor of 2 (Backeberg et al., 2014). Further, excessive EKE south of Madagascar and east of the Agulhas Current, present in the un-assimilated HYCOM configuration appears reduced, though the latter remains too high compared to the drifter data (Backeberg et al., 2014). This lingering excessive EKE is attributed to the consistent over-stimulation of eddy propagation along a common pathway from the main current to the retroflection region (Backeberg et al., 2014). The assimilated configuration's distribution of EKE near the retroflection region, as well as the way it is directed into the South Atlantic ocean are also shown to be improved. However, little improvement in the under representation of EKE associated with the Agulhas Return current, which is evident in the un-

assimilated configuration, is achieved by the assimilated configuration. It remains too weak when compared to that computed from the observational data. Findings also suggest that satellite altimetry fails, in some cases, to resolve sufficiently high levels of EKE in regions of high mesoscale variability (Backeberg et al., 2014). This is an important caveat in the use of altimetry as a reference when evaluating model simulations. Summarily, it is suggested that the improvements of the assimilated configuration are focussed in regions characterized by high levels of mesoscale variability. These are the Mozambique Channel, two regions to the south and south west of Madagascar and near the Agulhas retroflection (Backeberg et al., 2014).

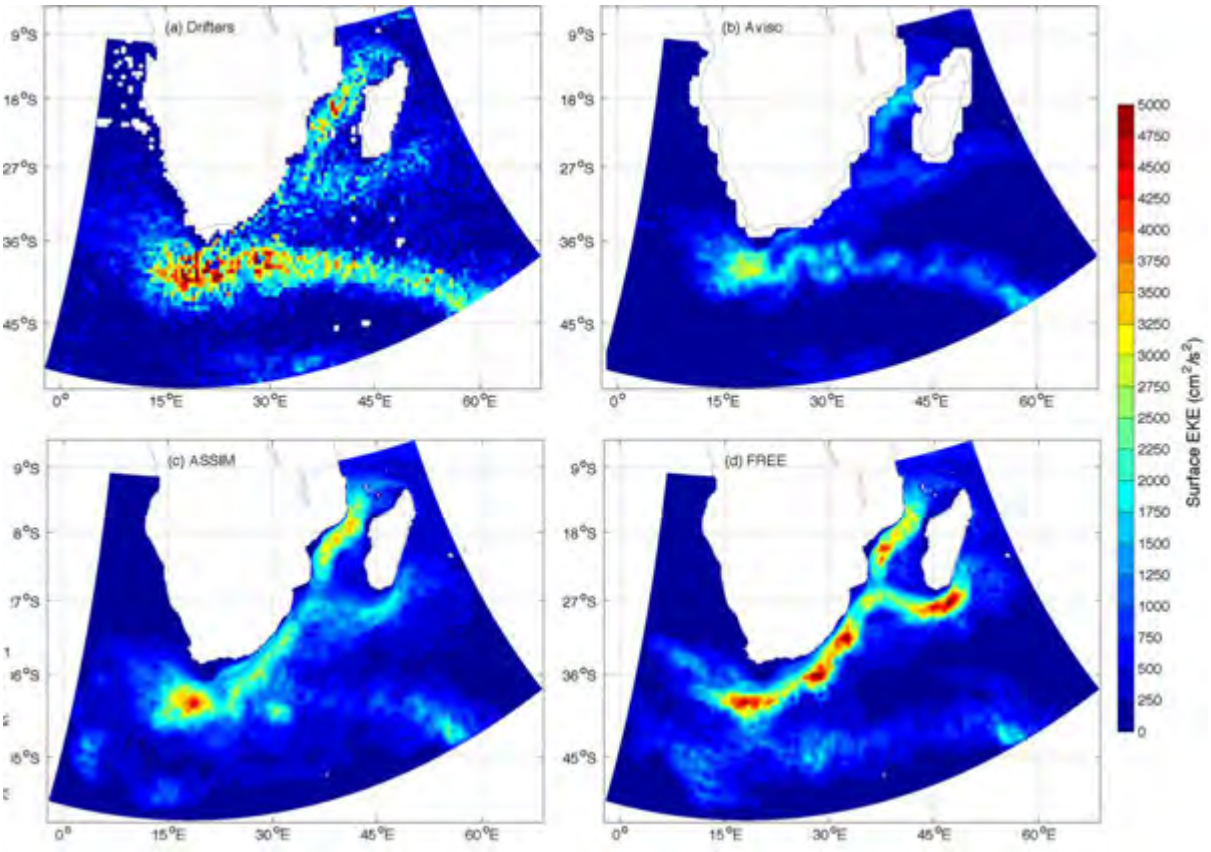

Figure 2. Mean surface eddy kinetic energy (cm2/s2) derived from surface velocities of (a) the drifter data, (b) Aviso NRT persistence, (c) Assim. and (d) Free. Figure taken from Backeberg et al., 2014.

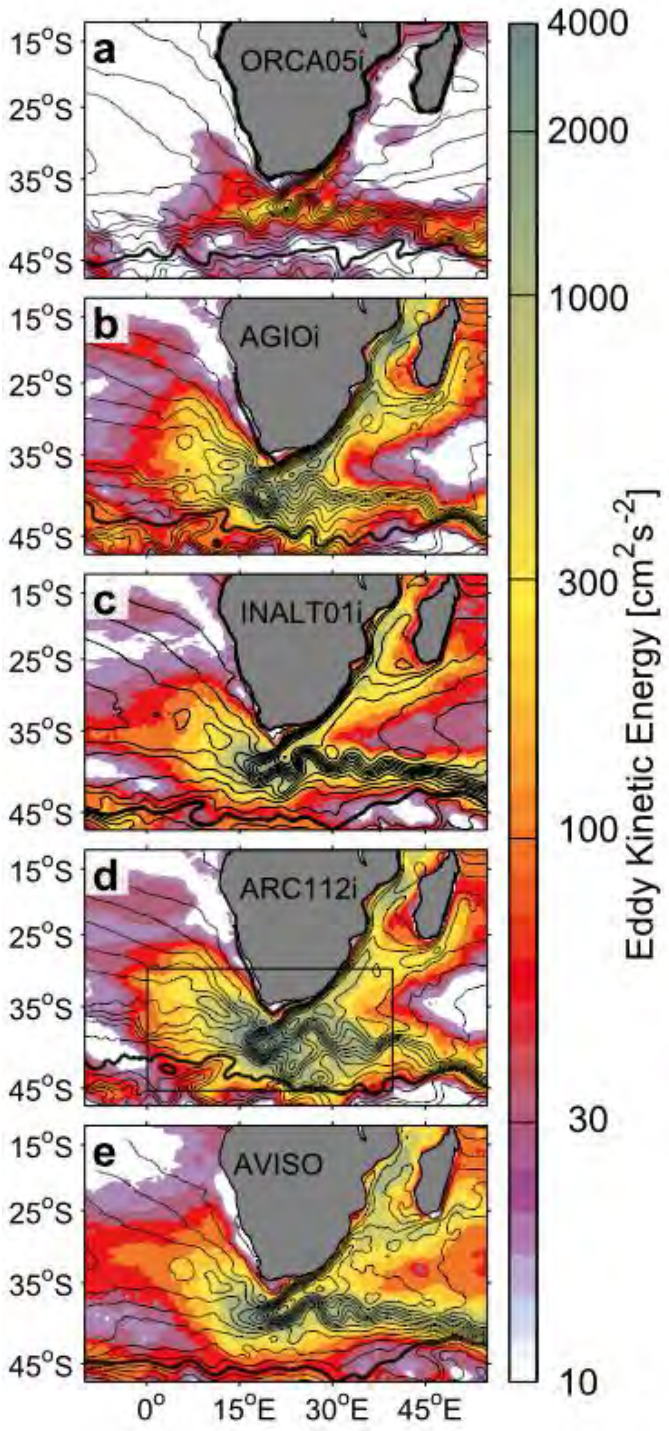

Figure 3. EKE for 1992–2007 from (a) ORCA05i, (b) AGIOi, (c) INALT01i, (d) ARC112i, and (e) Aviso. EKE is calculated from geostrophic velocities derived from the SSH. Contours show 10-cm delineations of mean SSH for this period. Figure taken from Loveday et al., 2014.

Loveday et al. (2014) assess four sets of model results in order to robustly quantify Agulhas leakage. These include two Regional Ocean Modelling System (ROMS) configurations (AGIOi and ARC112i) and two Nucleus for European Modelling of the Ocean (NEMO) configurations (ORCA05i and INALT01i). High values of EKE in the Mozambique Channel, as represented by AGIOi, ARC112i, and INALT01i, are highlighted as showing good agreement with observation (Loveday et al., 2014). Contrastingly, EKE south of Madagascar appears underrepresented by the models with respect to observation. Near the retroflection, the spatial distribution of EKE simulated by the models agrees with that of observation, and is associated with the propagation of Agulhas rings into the south Atlantic (Loveday et al., 2014).

Halo et al. (2014a) use model outputs from configurations of the Regional Ocean Modelling System (ROMS), the Hybrid Coordinate Ocean Model (HYCOM), Ssalto/Duacs satellite altimetry data and in-situ data from the Long-Term Ocean Climate Observation programme (LOCO) (de Ruijter et al., 2006; Ridderinkhof et al., 2010), to describe and compare eddy properties in the Mozambique Channel. The study employs an automatic eddy detection and tracking algorithm

to identify eddies (Halo et al., 2014a). Only eddies with lifetimes of greater than 30 days are tracked. A clearly defined southward propagation, near to the Mozambique coast, is evident among anticyclones. Cyclones exhibited a predominant south-westerly propagation (Halo et al., 2014a). Cyclones outnumber anticyclones in the Channel in all products and in general, exhibit shorter lifespans, lower amplitudes and smaller scales (Halo et al., 2014a). The total number of tracked eddies in this particular implementation of HYCOM is higher than that tracked in the Aviso altimetry data, with cyclones being 34.4 % and anticyclones 22.2 % more numerous than those of the observations (Halo et al., 2014a). Further, whilst HYCOM anticyclones appear reasonably sized, cyclones appear 23 % smaller than in other products. Aviso and HYCOM cyclones show normally distributed radii, centred around 70 km and 40 km respectively. Aviso and HYCOM anticyclones show bi-modally distributed radii, peaking at 30 km and 50 km in both products (Halo et al., 2014a). A similar trend is evident in amplitude, with cyclones exhibiting amplitudes some 45 % lower than those present in the altimetry. HYCOM cyclone lifetimes, however, are in better agreement with altimetry, despite anticyclones having lifetimes which are overestimated by around 22 % (Halo et al., 2014a).

Eddy properties derived from tracked eddies with a lifetime > 30 days, for number of eddies (N), mean lifetime ($\bar{\tau}$), mean amplitude ($\bar{\eta}$), and mean diameter ($\bar{L}$).

Eddy	N [eddies yr^{-1}]	$\bar{\tau}$ [day]	$\bar{\eta}$ [cm]	$\bar{L}$ [km]
Altimetry				
Cyclones	22.1	85.0	11.0	139.0
Anticyclones	17.1	101.0	14.0	157.0
HYCOM				
Cyclones	29.7	76.0	6.0	106.0
Anticyclones	20.9	123.0	12.0	138.0

Table 2. Eddy properties derived from tracked eddies with a lifetime > 30 days for number of eddies (N), mean lifetime ($\underline{\tau}$), mean amplitude ($\underline{\eta}$) and mean diameter ($\underline{L}$). Table adapted from Halo et al., 2014a).

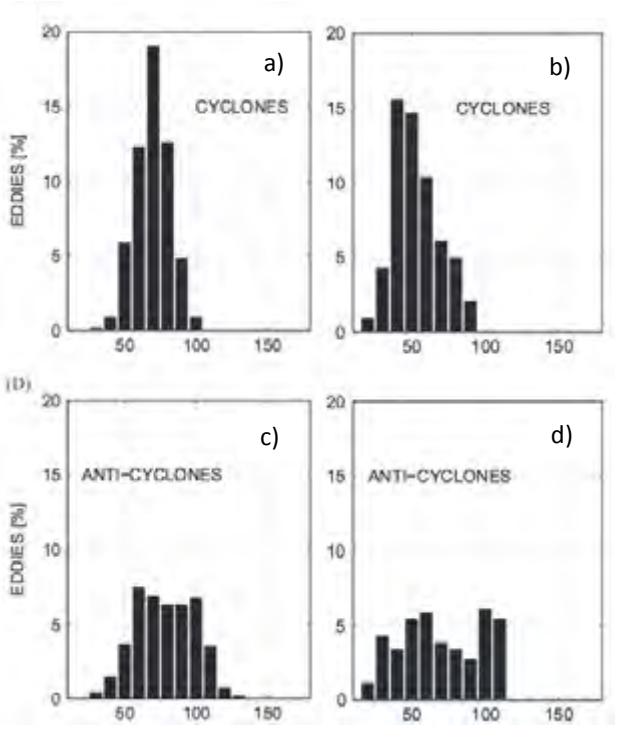

Figure 4. Eddy density distribution as a function of radius. (a & c), Aviso cyclones and anticyclones, (b & d) HYCOM cyclones and anticyclones. Figure adapted from Halo et al., 2014a.

Discrete analysis of eddies and their properties is less prevalent for the region south of Madagascar and the south-west Indian Ocean. Most recent studies tend to focus on the mesoscale dynamics associated with the Mozambique Channel, Agulhas Current system and south-east Atlantic Ocean. One study focussing on the region south of Madagascar is detailed in the work of de Ruijter et al. (2004), which investigated eddies and dipoles around south Madagascar. Observations made during the Dutch-South African Agulhas Current Sources Experiment (ACSEX) research cruise suggested that mesoscale, dipolar features, each consisting of adjacent cyclones and anticyclones, occur south and south-west of Madagascar. Drogues were deployed in order to track these features.

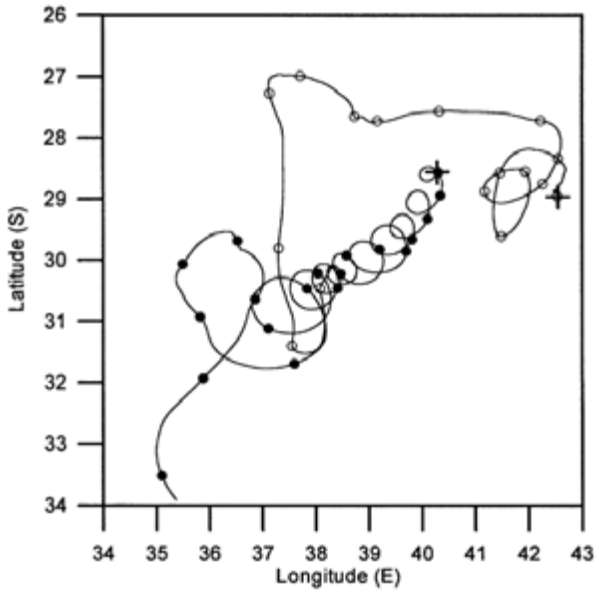

Figure 5. The tracks of two shallow drogued surface drifters, deployed in the cyclonic eddy (black symbols) and the anticyclonic eddy (open symbols) during the ACSEX cruise. The time interval between successive symbols is 5 days. The deployment positions are indicated with a cross. The position data have been smoothed in order to remove any tidal signal. Figure taken from de Ruijter et al., 2004.

Subsequent analysis of TOPEX-POSEIDON/ERS (TP/ERS) altimetry data revealed that these eddies might be linked to Indian Ocean Dipole and El Niño phases (which coincided with peak formation rates) (de Ruijter et al., 2004). Eddies were observed to behave fairly irregularly, often splitting and merging with other eddies (de Ruijter et al., 2004).

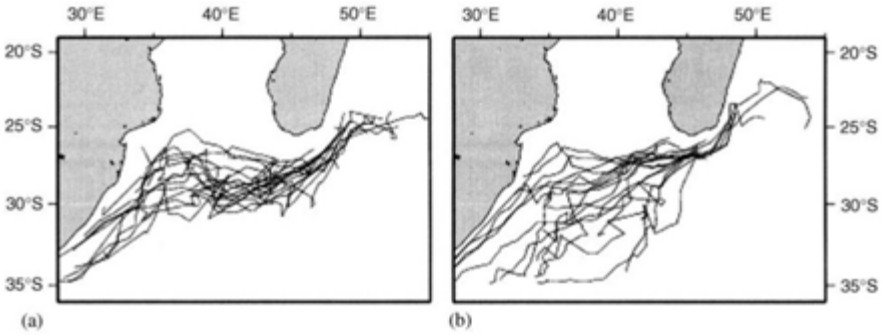

Figure 6. Tracks of the Madagascar anticyclones (a) and cyclones (b) after their first identification as eddy pairs in the period April 1995– December 2000 in the combined TP/ERS altimetry dataset. Figure taken from de Ruijter et al., 2004.

In total, 16 dipole pairs are identified and tracked visually in this region. Both those eddies tracked by the deployment of drogues, as well as those identified from altimetry data, exhibit a general westward (and sometimes north-westward), followed by south-westward drift in the case of anticyclones. Cyclones tend to drift south-westward, but show more divergence in their tracks.

2.2 Data Assimilation, its Suitability and Potential Impacts

Data assimilation is a strategic process aimed at improving the accuracy, efficiency and realism of a predictive model solution (Robinson & Lermusiaux, 2000). It can also be thought of as the production of initial conditions via a statistical combination of observations and short-range forecasts (Kalnay, 2003). Talagrand (1997) simplifies and broadens this idea, in describing data assimilation as the use of all available information in the determination of the possible state of the atmospheric or oceanic flow. In order to do this, it combines observational data with the underlying dynamical principles which govern the system in question (in this case, the ocean). Data assimilation has enabled rapid and important advances to be made in ocean science. (Robinson & Lermusiaux, 2000). Acknowledging that ocean prediction is an initial-value problem (given the state of the ocean at a certain time, determine its state at a later time) (Robinson & Lermusiaux, 2000), the capacity to establish accurate initial conditions for a prediction system is particularly important, and a problem with which data assimilation can help (Kalnay, 2003).

The general process of estimating state variables (e.g. velocity components, pressure, density, temperature) and parameters (e.g. diffusivities, rates of earth rotation, viscosities) involves the linking of state variables to sensor or initial data. Model dynamics interpolate and extrapolate the data, and linkages between the dynamics allow all state variables and parameters in question to be estimated from measurements of some of them (those which can be obtained practically, in reality). The data are assessed for their errors, and combined with the dynamics such that those with larger errors carry smaller weightings. The system is generally nonlinear,

and subject to numerous feedbacks. (Robinson & Lermusiaux, 2000). Figure 7 provides a schematic visualisation of some of these interactions.

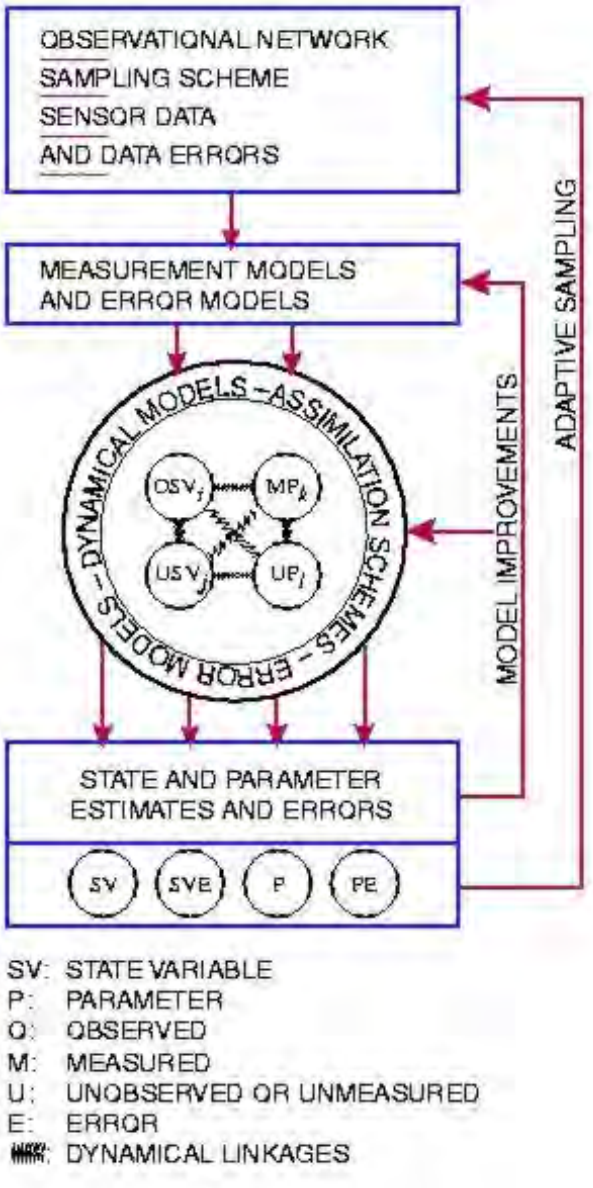

Figure 7. A schematic of the general process of data assimilation and state and parameter estimation. Figure taken from Robinson & Lermusiaux, 2000.

It is important to evaluate data assimilation efforts constantly in order to ensure that data assimilation does not "shock" a numerical solution. The evaluation of assimilated model outputs requires the systematic assessment of key parameters against climatological and/or real time data and quality controlled observations (Chassignet et al., 2007). The nudging of model solutions via the assimilation of observational data should not be so drastic as to cause an unrealistically large or rapid change in properties. Whilst model solutions can be artificially forced into agreement with observations, inherent improvements with respect to independent observational data are not trivial (Chassignet et al., 2007). Degradation of model solutions as a result of data assimilation is possible where assumptions used by the assimilation scheme are inaccurate (Chassignet et al., 2007). The selection of data, as well as the scheme via which they are assimilated should thus be strategically sound.

The success of data assimilation in improving model performance also depends on the appropriate management of practical implementation details. These might include state vector structure, correlation scales, initialization procedure, covariance rank and the assimilation of multivariate observation types (Srinivasan et al., 2011).

A further practical consideration in the effective use of data assimilation is computational cost. Whilst certain schemes might yield better accuracy, high computational costs may hinder the operational value of a data-assimilated configuration (Srinivasan et al., 2011).

In this regard, the EnOI is suitable for the context of this study, producing a reasonable computational cost (Backeberg et al., 2014).

Recently, the EnOI has been successfully applied in the assimilation of along-track SLA data into an eddy resolving model of the South China Sea. In that study, the Root Mean Square Error of sea level anomalies is reduced from 9.3 cm to 6.9 cm, and the overall representation of mesoscale features is improved (Xie et al., 2011). It has been implemented in further regions dynamically similar to the Agulhas domain; namely in the Gulf of Mexico (Counillon and Bertino, 2009a) and the Australian region (Oke et al., 2007).

In this way, data assimilation applied to HYCOM via the EnOI is a strategically sound approach to improving the simulation of mesoscale features.

Chapter 3. Methodology

3.1 Models, Data and Data Assimilation

A data-assimilative ocean prediction system generally requires a sufficiently robust and well-understood dynamical ocean model, a data assimilation scheme and recent observational data (Backeberg et al., 2014). This study uses along-track satellite sea-level anomaly data from Ssalto/Duacs, distributed by Aviso (observational data). These data are assimilated into a HYCOM configuration (dynamical model) using the Ensemble Optimal Interpolation (EnOI) (assimilation scheme) (Backeberg et al., 2014). HYCOM has been employed with reasonable success in the representation of important elements of the Agulhas system (Backeberg et al., 2008, 2009), whilst the EnOI assimilation scheme (Evensen, 2003; Oke et al., 2002) has been shown to be effective in regions with similar dynamics (Oke et al., 2005, 2007; Counillon & Bertino, 2009a, b; Xie et al., 2011; Srinivasan et al., 2011).

3.1.1 Hybrid Coordinate Ocean Model

The Hybrid Coordinate Ocean Model (HYCOM) is a primitive equation ocean model which uses a pressure-density hybrid vertical coordinate system derived from the Miami Isopycnal Coordinate Model (MICOM) (Griffies et al., 2000). Given the importance of vertical coordinate system selection for modelling studies (Griffies et al., 2000), HYCOM employs a combination of isopycnic (ρ), terrain-following (σ) and Cartesian (z) vertical coordinate systems such that the vertical layer discretization adapts to optimize its structure depending on the dominant processes in a particular layer for each time-step (Bleck, 2002). A ρ coordinate system is used for the open ocean, z-level coordinate system for the mixed layer and a σ coordinate system for shallow, inshore regions where bathymetry is influential (Bleck, 2002). This study employs a nested HYCOM configuration within a basin-scale parent grid (George et al., 2010). The nested configuration receives snapshot fields as boundary conditions from the parent grid on a 6-hourly basis. Certain slow-varying parameters such as temperature, salinity layer interface and baroclinic velocities are relaxed towards the parent grid over a distance of 20 grid-cells (Backeberg et al., 2014). The nested domain extends from 10°S - 50°S, and from 0° to 60°E in order to provide regional coverage of the Agulhas system. The horizontal resolution of the parent grid ranges from 14 km in the north to 45 km in the south. The nested grid resolution is 1/10°, or 10 km. Given a regional average Rossby radius of deformation of 30 km (Chelton et al., 1998), the nested grid is considered eddy-resolving. This configuration allows for the ratio of Parent: Nest grid resolution to be kept within the recommended range of 3:1 - 5:1, an important consideration in ensuring the stability of the model (Counillon & Lisæter, 2011).

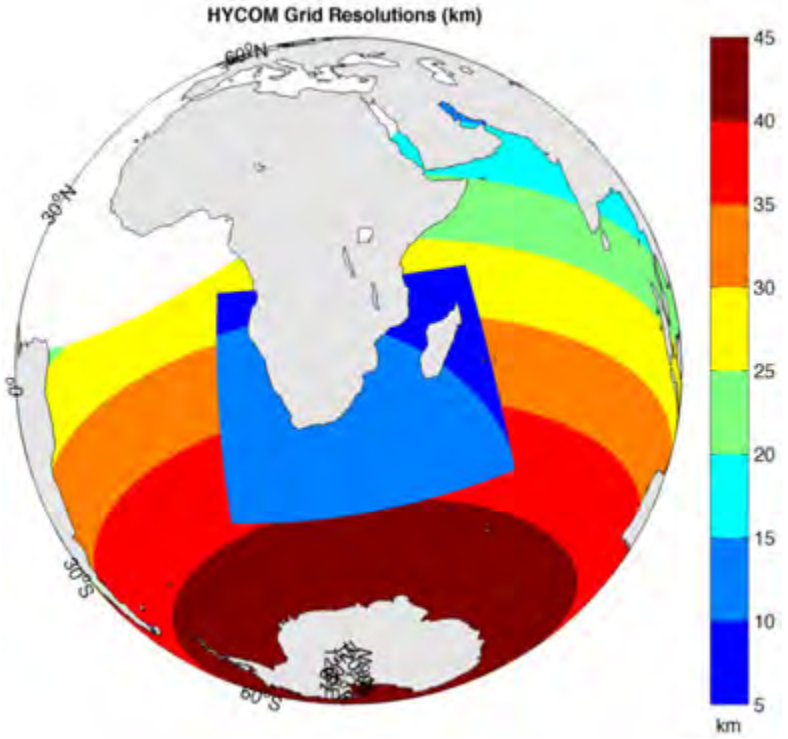

Figure 8. Latitudinal variation in horizontal grid resolution. The coarser resolution of the basin-scale, parent grid is evident in the zonal colour bands. The finer resolution of the nested grid is visible in the two darkest shades of blue (5-10 km and 10-15 km resolution ranges).

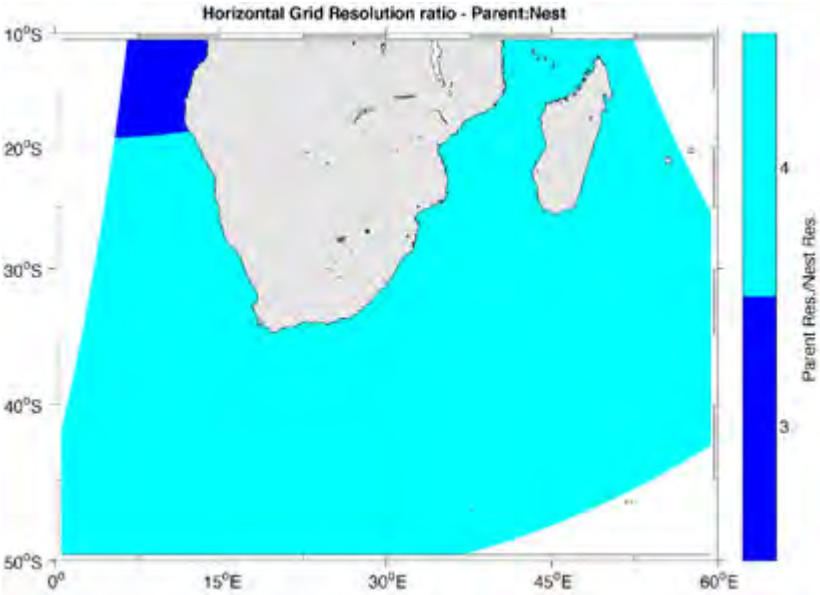

Figure 9. Horizontal grid resolution ratio between the parent and nested grid (Parent: Nest). This figure shows the acceptable ratio (> 5:1) achieved throughout the domain between the parent and nested grid resolutions; an important consideration for the stability of nested configurations.

3.1.2 Satellite Observations

Satellite altimetry data for assimilation into HYCOM were produced by Ssalto/Duacs and distributed by Aviso. This is a composite, delayed time unfiltered SLA dataset, comprised of altimetry data from four satellites; Jason-1 (January-October 2008), Jason-2 (October 2008 onwards), Jason-1 new orbit (February 2009 onwards), Envisat and GFO (January-September 2008). Support for this product is provided by CNES (http://www.aviso.oceanobs.com/duacs/)

(Backeberg et al., 2014). Delayed time, unfiltered along-track SLA data are assimilated in order to avoid problems associated with the interpolative smoothing of the gridded product (Ducet et al., 2009) which arise in regions of rapid change (Backeberg et al., 2014). The horizontal resolution of the unfiltered data is 7 km. For the study period (2008-2009), data from Jason-1 (start-October 2008), Jason-2 (October 2008-February 2009), Jason-1 in its new orbit (February 2009-end), Envisat and GFO were incorporated into the composite product (Backeberg et al., 2014).

The satellite data used for comparison with HYCOM in this study are generated in the same way as those used for assimilation (courtesy of Ssalto/Duacs and Aviso), but are packaged as gridded products with a resolution of 0.25°. For eddy detection and tracking, the algorithm is applied to the merged Global Ocean Gridded sea Absolute Dynamic Topography SSALTO/Duacs L4 product. It contains Delayed Time Level-4 sea surface height above Geoid products from multi-satellite observations over Global Ocean. For easier visualization, the corresponding SLA data are used in some analyses, as eddies detected by the algorithm (see 3.2) are more clearly defined on maps of SLA than ADT. The delayed Time merged Global Ocean Gridded Sea Level Anomalies SSALTO/Duacs L4 product was selected for this purpose, with SLA referenced to the 1993-2012 period.

3.1.3 EnOI

Optimal interpolation data assimilation methods have been a prominent choice for operational analysis since the late 1980s (Kalnay, 2003). The Ensemble Optimal Interpolation (EnOI; Oke et al., 2002; Evenson et al., 2003), a derivative of the Ensemble Kalman Filter (EnKF) is selected as the data assimilation scheme for the HYCOM simulation. The implementation of this scheme in this context is as described in Backeberg et al. (2014). It employs a static ensemble ($A \in R^{n \times N}$) where N is the ensemble size and n is the model state dimension (Backeberg et al., 2014). The EnOI assumes the climatological variability to be representative of the forecast error developed during an assimilation cycle, unlike the EnKF which uses Monte Carlo methodology to construct the flow-dependent covariance C (Backeberg et al., 2014):

Equation 1.

$$\overline{\varepsilon\varepsilon^T} \approx C = \frac{\alpha}{N-1} A'(A')^T$$

where $A' = A - \overline{A}$ (i.e. A' is the centred historical ensemble with A as the historical ensemble comprised of model states) (Counillon & Bertino, 2009a). α Is used to moderate the ensemble variance constantly (Backeberg et al., 2014).

The classical Kalman Filter equation can be solved as follows:

Equation 2. $$\psi^a = \psi^f + \alpha A' A'^T H^T (\alpha H A' A'^T H^T + TT^T)^{-1} (d - H\psi^f)$$

Where the estimate ψ^a is a combination of *centred model states* from the ensemble and the *forecast* ψ^f (Counillon & Bertino, 2009a). d denotes the measurements (assimilated data), H is the measurement operator which relates the state variables of the model to the measurements.

3.2 Eddy Detection and Tracking Algorithm

The automated eddy detection algorithm implemented in this study employs a combination of two methods; geometric and dynamical. It is coupled with a tracking scheme and coded in MATLAB by Dr. Pierrick Penven. The package is freely available for download at http://www.simocean.org.za/tooleddy.php. This tool provides an objective method for the identification and tracking of important regional mesoscale features, as it has minimal tunable parameters and can be applied to various different data products. This kind of consistency is critical for the robust comparison and quantification of eddy properties as simulated by different models and observations (Halo et al., 2014a).

The geometric detection method is typical of those applied to altimetry data (Halo et al., 2014a). It identifies local minima and maxima in SSH fields as eddy-centres, and the largest corresponding closed-loop of SSH (Chelton et al., 2011) as eddy boundaries. In this case, local maxima (minima) are identified as grid points with higher (lower) SSH than their eight surrounding grid points.

The dynamical method seeks to classify the kind of flow associated with a feature. If the modification of the flow is dominated by vorticity rather than strain (Souza et al., 2011), the feature is classified as an eddy. Conversely, if the modification of the flow is due mostly to shear and normal components, the feature is not considered an eddy (Halo et al., 2014a). This kind of assessment demands some reference, and to this end the Okubo-Weiss parameter (Okubo, 1970; Weiss, 1991) is computed and assessed against a threshold value. The Okubo-Weiss parameter is defined as:

Equation 3. $$W = S_n^2 + S_s^2 - \xi^2$$

where

Equation 4. $$S_n = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}$$

Equation 5. $$S_s = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

Equation 6. $$\xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

S_n is the shear component and S_s the normal component of the strain tensor, whilst ξ is the vertical component of relative vorticity. u and v are the zonal (in x) and meridional (in y) components of geostrophic velocity, respectively (Halo et al., 2014b). They are computed from the SSH field as follows:

Equation 7. $$\mathrm{u} = \frac{-\mathrm{g}\,\partial[\mathrm{SSH}]}{\mathrm{f}\,\partial \mathrm{y}}$$

Equation 8. $$\mathrm{v} = \frac{\mathrm{g}\,\partial[\mathrm{SSH}]}{\mathrm{f}\,\partial \mathrm{x}}$$

where g is the gravitational acceleration and f is the Coriolis parameter. A threshold of $W = 0$ is set, such that where $W < 0$, the flow is deemed to be dominated by vorticity and an eddy exists. This threshold is one of the few explicit tunable parameters of the detection scheme, and has been set as $W = 0$ in accordance with numerous previous studies (Pasquero et al., 2001; Isern-Fontanet et al., 2006; Henson & Thomas, 2008; Souza et al., 2011; Halo et al., 2014a; Halo et al., 2014b).

The package employed in this study combines the geometric and dynamical detection criteria, such that a feature incorporating of closed loops of SSH *and* a negative Okubo-Weiss parameter is considered an eddy. This combined method has been shown to be more robust in the consistent detection of eddies (Halo et al., 2014b). One further condition for eddy detection is a radius of less than 400 km. This is to avoid basin-scale ocean gyres from being detected as mesoscale eddies (Halo et al., 2014b). Applying the algorithm to absolute SSH, as opposed to SLA data, provides two main benefits. Firstly, it avoids the detection of current meanders as mesoscale eddies (Halo et al., 2014a). Secondly, it avoids the recognition of negative SSH anomalies as cyclonic eddies, rather than mere absences of anticyclonic eddies. This is important in a region where large, anticyclonic eddies are numerous (de Ruijter et al., 2002).

Previous applications of this particular detection and tracking algorithm define a constant contour interval for the detection of closed loops of SSH. This is another of the few tunable parameters on which the algorithm depends. In this study, the algorithm was first used with a set contour interval of 2 cm, based on previous studies and their assessment of the average precision of Aviso altimetry (Halo et al., 2014a). Though it has been shown that the algorithm is

not particularly sensitive to the definition of this parameter (Halo et al., 2014a), concerns over the range in precision of the altimetry over the specific study domain led to the incorporation of a spatially-varying contour interval, according to the expected precision of the altimetry data for a given location. This was achieved by modifying the algorithm code such that it draws on the observational *representivity error* for a given location, rather than a set contour interval throughout the entire domain. Representivity error arises from inconsistencies between the model and observational space due to, for example, different resolutions. In this case, representivity error is computed from the SSH variance derived from gridded satellite SSH, scaled to 2.

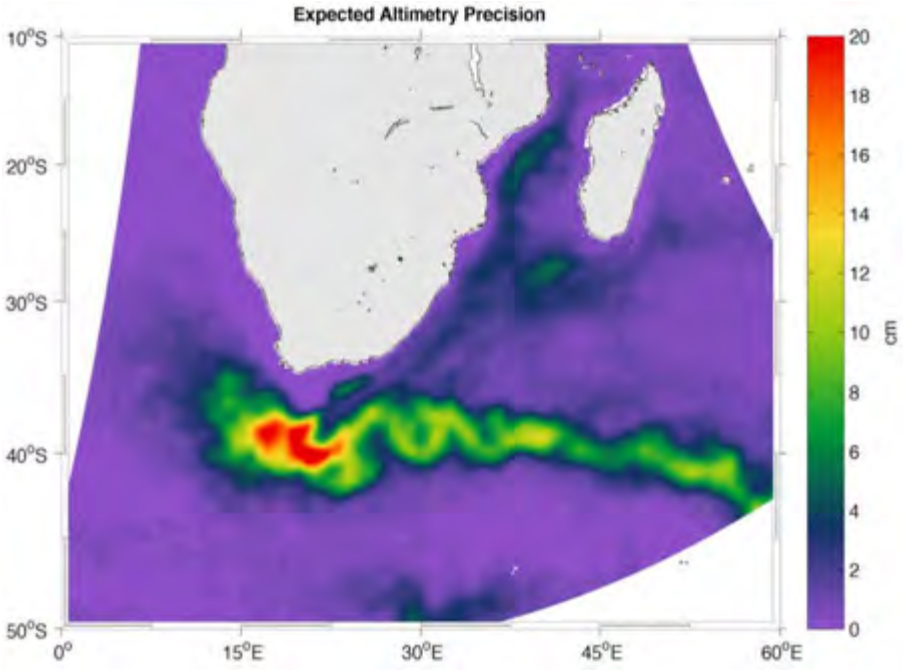

Figure 10. Expected precision (representivity error) of Aviso satellite altimetry data. This spatially-varying parameter is incorporated into the automatic eddy-detection algorithm by modifying the algorithm code for use in this study.

The tracking of eddies is accomplished via the methodology of Penven et al. (2005), as implemented by Halo et al. (2014a; 2014b). It strives to assess the similarity of eddies in subsequent timesteps. In subsequent frames, two eddies with minimum generalised change to some of their properties in a non-dimensional property space are identified as the same eddy (Halo et al., 2014a).

Equation 9.

$$X_{e1,e2} = \sqrt{\left(\frac{\Delta X}{X_0}\right)^2 + \left(\frac{\Delta R}{R_0}\right)^2 + \left(\frac{\Delta \xi}{\xi_0}\right)^2}$$

where $\Delta X$is the distance between two eddy centres e_1 and e_2, $\Delta R$is the change in eddy diameter, $\Delta\xi$ is the change in eddy vorticity, X_0 is a typical length scale (in this case, 25 km as per Penven et al. (2005); Halo et al. (2014a,b)) $R_0$is a typical radius (in this case, 200 km) and ξ_0 is a typical vorticity (in this case, $10^{-5}.s^{-1}$. In order to prevent anticyclones becoming cyclones and vice-versa, if a propagation of $\Delta X$is accompanied by a change in sign of vorticity, $X_{e1,\ e2}$ is set to infinity (Halo et al., 2014a).

Furthermore, a threshold is defined such that eddies exhibiting displacements implying unrealistically high drift-velocities between subsequent frames cannot be the same eddies:

Equation 10.
$$\mathrm{U_{max}} = 0.3\ \mathrm{m.s^{-1}}$$

where U_{max} is the maximum permissible derived drift speed between frames.

Summarily, the criteria for detection and tracking are as follows:

Detection: A feature is an eddy if it exhibits:

- At least one closed contour of SSH ***and***
- An Okubo-Weiss parameter < 0.
- A radius < 400 km.

Tracking: An eddy retains its identity in subsequent frames if:

- The change in some of its properties is minimum ***and***
- Implied drift speed < 0.3 m.s^{-1}.

3.3 Computation of Eddy Properties

3.3.1 Area

Eddy area is computed from spheric distances between the relevant latitudes and longitudes of grid cells containing an eddy.

3.3.2 Eddy Radius

Eddy radii are computed according to:

Equation 11.
$$\mathrm{R} = \sqrt{\frac{\mathrm{A}}{\pi}}$$

where *A* is the Area as per 3.3.1.

3.3.3 Vorticity

Eddy relative vorticity is computed from the geostrophic flow components (u, v as derived in the Okubo-Weiss calculation):

Equation 12. $$\xi = \frac{\partial \mathrm{v}}{\partial \mathrm{x}} - \frac{\partial \mathrm{u}}{\partial \mathrm{y}}$$

The result is then averaged over the eddy radius to give a mean relative vorticity for a particular eddy (Penven et al., 2005).

3.3.4 Amplitude

Eddy amplitude is computed as the difference between maximum SSH and minimum SSH for the area encompassed by an eddy (within the boundary contour).

Equation 13. $$\mathrm{a} = \mathrm{SSH}_{\max} - \mathrm{SSH}_{\min}$$

3.3.5 Eddy Kinetic Energy

As with *eddy vorticity*, EKE is computed from derived eddy geostrophic flow components:

Equation 14. $$\mathrm{EKE} = \frac{\mathrm{u}^2 + \mathrm{v}^2}{2}$$

where u is the zonal flow component and v is the meridional flow component as calculated in Equations 7 & 8 respectively.

3.3.6 Rotational Speed

Eddy rotational speed is computed as the maximum average geostrophic speed (calculated as per section 3.2) within an eddy for a range of SSH (from the centre to the eddy boundary). This is the speed at which a particle trapped within the eddy might travel and follows the methodology of Chelton et al. (2011).

3.3.7 Translational Speed

Eddy propagation or drift speed components are computed by the eddy tracking algorithm:

Equation 15. $$\mathrm{U} = \frac{\partial \mathrm{x}}{\partial \mathrm{t}}$$

Equation 16. $$\mathrm{V} = \frac{\partial \mathrm{y}}{\partial \mathrm{t}}$$

where U is zonal propagation speed, V is meridional propagation speed, and δx,δy and δt are the zonal and meridional changes in eddy centre position and change in time, respectively, between frames.

3.3.8 Advective Nonlinearity

Eddy nonlinearity is computed from propagation or translational speed and rotational speed:

Equation 17.
$$\text{Non Linearity} = \frac{\mathrm{U}}{\mathrm{c}}$$

where

Equation 18.
$$\mathrm{c} = \sqrt{\mathrm{u}^2 + \mathrm{v}^2}$$

where U is rotational speed as derived previously and c is translational speed, derived from geostrophic flow components as before.

3.4 Preparation of Data for Detection & Tracking

In order to ensure that the detection and tracking scheme is given the same chance to detect eddies in three different datasets, comparability of the three datasets needs to be checked. Several steps are taken in this regard. Firstly, the daily Aviso data are averaged to provide weekly mean SSH fields coinciding temporally with the HYCOM weekly mean SSH fields. Secondly, satellite data (gridded by default onto a regular Mercator projection) falling outside the curvilinear HYCOM grid is disregarded. Finally, a buffer zone of 30 grid cells is defined along all boundaries, and the detection and tracking scheme applied to data within this buffer zone for all three datasets. This is to avoid consideration of any spurious features which might be the result of near-boundary model inaccuracies.

3.5 Gridded Analysis

Results from the eddy detection and tracking scheme are in vector form. However, sample sizes (numbers of eddies) and their spatial distribution vary between datasets (models and altimetry) and so a second approach was sought in order to achieve spatially coherent comparisons. In order to facilitate easier interpretation, visualization and robust analysis, results are gridded and mapped over the study domain. After starting trials with higher resolutions, the analysis grid resolution was systematically degraded until a resolution of 1° x 1° was selected (except in the case of *density distribution*, Figure 13, which was kept at 2 °). This reduces the amount of noise when mapping the various fields, and allows for better continuity between grid cells, aiding interpretation.

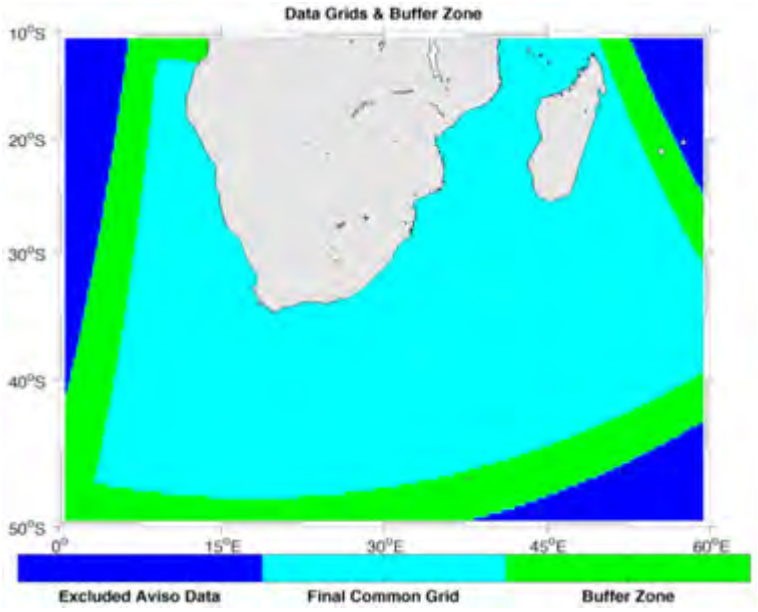

Figure 11. Adaptation of the different grids. Aviso data are gridded by default onto a grid which would cover the entire study domain; it would include data in the dark blue regions. The boundaries of the curvilinear HYCOM grid are visible as the outer boundaries of the green region. 30 grid cells inwards from these boundaries are defined as a buffer zone, to give the final, common grid (cyan) to which the detection and tracking scheme is applied.

3.6 Error Calculations

In order to quantify the effect of data assimilation, the RMSE for the vector results (those generated directly from the detection and tracking scheme) is computed.

Equation 19.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(y_o - y_e)^2}{n}}$$

where n is the number of estimations $y_o$is the truth or reference value (derived here from satellite altimetry observations), and $y_e$is the estimated value (derived from the model SSH).

The difference in percentage error is then calculated for each grid cell where eddies were found during the study period, with $y_o$being the mean true (satellite) value for that property in that grid cell and $y_e$being the mean estimated (model) value for that property in the same grid cell.

Equation 20.

$$\%\ \mathrm{Error} = \frac{y_e - y_o}{|y_o|} \times 100$$

The choice of gridded Aviso SSH product as the reference is a pragmatic one (satellite altimetry being the only product to provide spatially and temporally sufficient SSH observations for use as a control).

The final step in error analysis is to compute the difference in error of the two model simulations (Free and Assimilated). This is achieved by computing the difference in percentage error for each grid cell for a particular eddy property. The difference in percentage error is then mapped. This has the advantage of providing insight into the spatial variation of percentage

error of the products in simulating eddies, which is important when a product may simulate features well in one location but poorly in another. Chassignet et al. (2007) note that the evaluation of data-assimilated model outputs against unassimilated, independent measurements is an effective strategy in assessing performance of those outputs. The differences between models and observations determined in this way serve as common indicators of accuracy (Chassignet et al., 2007).

Equation 21. $$\text{Relative Error} = \left|\%\ \text{Error}_{free} - \%\ \text{Error}_{assim}\right|$$

Implicitly, this kind of relative error mapping is only valid where all three products produce at least some eddies during the two years. Relative error cannot be computed where the two simulations and the altimetry did not all produce eddies, as for those grid cells, there would be no eddy properties to compare. Thus, the error mapping technique gives information about the general accuracy with which Free and Assim. simulate eddies in areas where both runs and the altimetry produced eddies. It cannot give information about the accuracy of eddy characteristics in areas where, for example, Assim. run produced eddies, but Free or Aviso data did not. This is useful insofar as comparisons of eddy properties are only made in areas where eddies actually existed in both simulations and in the observations. This is more robust than computing and comparing eddy properties (e.g. eddy kinetic energy from surface current velocities) from continuous variable fields where actual eddies may not exist.

Chapter 4. Results & Discussion

4.1 Sample Totals

	No. Cyclones		No. Anticyclones		Total Eddy Count		Differences	
	2008	**2009**	**2008**	**2009**	**2008**	**2009**	**2008**	**2009**
Free	560	569	549	512	1109	1081	-383	-361
Assim.	919	933	885	927	1804	1860	+312	+418
Aviso	767	731	725	711	1492	1442		

Table 3. Total count of mesoscale eddies with lifetimes of 7 days or more for the entire study domain, 2008/9. Differences are computed with respect to the corresponding number of cyclones and anticyclones detected by the altimetry data. A positive (negative) difference indicates more (fewer) eddies than those produced by Aviso. This provides a broad overview of the under-stimulation of mesoscale features by Free simulation, and the converse for Assim. simulation.

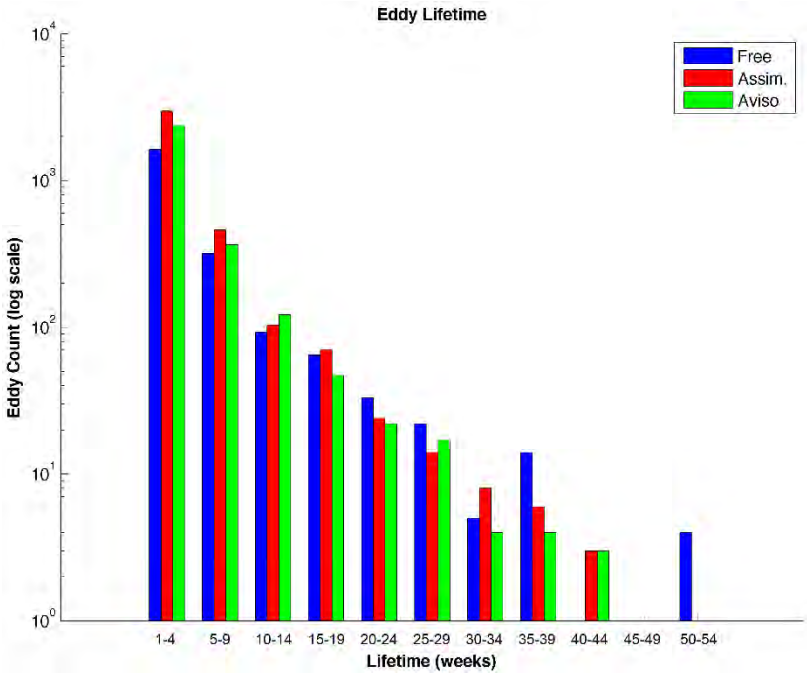

Figure 12. Distribution of eddy lifetimes for eddies lasting at least 7 days (1 week). Results are highly similar for cyclones and anticyclones. Eddy count is presented on a logarithmic scale along the vertical axis. The vast majority of eddies resolved by all three products have lifetimes ranging from 0-5 weeks. 75 % of Free eddies, 81 % of Assim. eddies and 80 % of Aviso eddies have lifetimes in this range. Good general agreement in eddy lifetime distribution is observed between all three products.

The sample totals given in Table 3 indicate similar trends to those seen in the altimetry and HYCOM results of Halo et al., (2014a), with cyclones outnumbering anticyclones in all three products. However, the differences between numbers of cyclones and anticyclones are much smaller in this study. Halo et al. (2014a) found cyclones to be, on average, 12.8 % and 17.4 % more numerous than anticyclones in the altimetry and HYCOM respectively (Halo et al., 2014a). In this study, cyclones are, on average, 3.1 %, 1.1 % and 2.12 % more numerous than anticyclones in Free, Assim. and Aviso respectively for the two-year period. Interestingly, however, Free simulated fewer eddies than those detected in the altimetry data, whilst Halo et al. (2014a) found the free-running HYCOM simulation to simulate more eddies than the altimetry. Contrastingly, Assim. simulates considerably more eddies than those detected in the altimetry.

Agreement is less clear with the results of Richardson (2007). Drifter data from that study indicate that anticyclones outnumber cyclones by 17.08 %. RAFOS floats show cyclones to outnumber anticyclones by 9.68 %, whilst the ALACE floats describe 3 anticyclones but do not provide any cyclone data. It should be borne in mind that the Richardson (2007) domain covers only that part of the Agulhas system extending westwards of 25 °E, within a similar latitudinal range (20 °S-45 °S) to this study.

The work of de Ruijter et al. (2004) is concerned with mesoscale dipole pairs, each consisting of a cyclone and an anticyclone, and thus tracks equal numbers of both features (de Ruijter et al., 2004).

Important caveats to bear in mind are that Halo et al. (2014a) focus almost exclusively on eddies found in the Mozambique Channel. There may be dynamical differences between that domain and the one in question, which might alter eddy distributions or, for example, the proportion of cyclones to anticyclones. Furthermore, analysis in that study is restricted to eddies living for a minimum of 30 days, whereas in this study, eddies of all lifetimes are considered. Secondly, the drifter and float data gathered by Richardson (2007) may provide good hydrographic in-situ data for the description of individual eddies, but suffers poor temporal and spatial resolution. Whilst they describe a similar meridional domain (20 ˚S – 45 ˚S), their zonal domain is much small than that of this study, extending from 10 ˚W to 20 ˚ E (compared to 0 ˚ - 50 ˚ E in this study). A cumulative total of 68 eddies between the drifter data (1994 – 2004), RAFOS float data (1997 – 1999) and ALACE float data (1996 – 2000) (Richardson, 2007) indicates that only a tiny fraction of the likely total number of eddies is captured. Further, eddies are detected and tracked visually, introducing some degree of subjectivity. As such, this study provides relatively good coverage of an otherwise sparse sphere of research.

4.2 Density Distribution

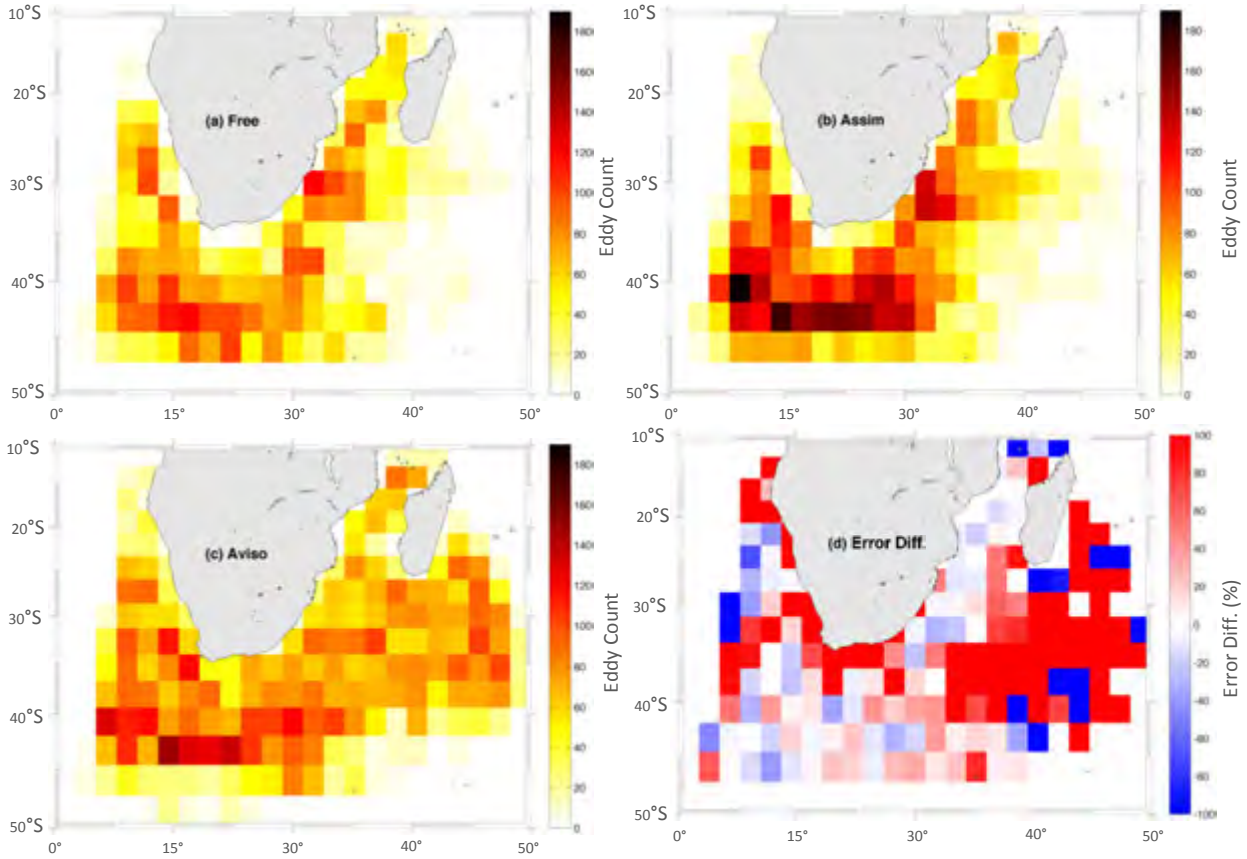

Figure 13. Spatial density distribution of eddies for 2008/9 for a) Free, b) Assim. and c) Aviso. Tile d) shows relative error; positive (negative) percentage values indicate closer similarity of Assim. (Free) to the altimetry. Red therefore represents improvement due to data assimilation, whilst blue represents degradation.

The assimilated run offers an improved overall eddy spatial distribution field over the free run. Though the generation of eddies remains insufficient in the region south of Madagascar and in the south-west Indian Ocean, the Assim. field shows more mesoscale activity than that of Free. The degree of improvement shown in the region south of Madagascar by the relative error analysis (tile *(d)*) seems incongruously high, given the lingering deficiency of eddies. This is likely due to the discretization of the calculation. Considering that Free has almost no eddies present in the grid cells in question, the presence of even a small number of eddies in the corresponding Assim. grid cells would yield vastly lower relative error. Good improvement is evident south (as far as ~ 40 ˚S) of the southern and south-western Cape coasts. South of 40 ˚S however, Assim. appears to over-stimulate eddies. This is likely the reason why, despite still under-stimulating eddies south of Madagascar and in the south-west Indian Ocean, the overall eddy count for Assim. is higher than for Aviso.

This over-stimulation is further apparent in the Mozambique Channel, where the number of simulated eddies remains too high. Though this is also the case in Free, it appears that the problem is slightly exacerbated by data assimilation.

4.3 Eddy Formation/Dissipation Sites & Trajectories

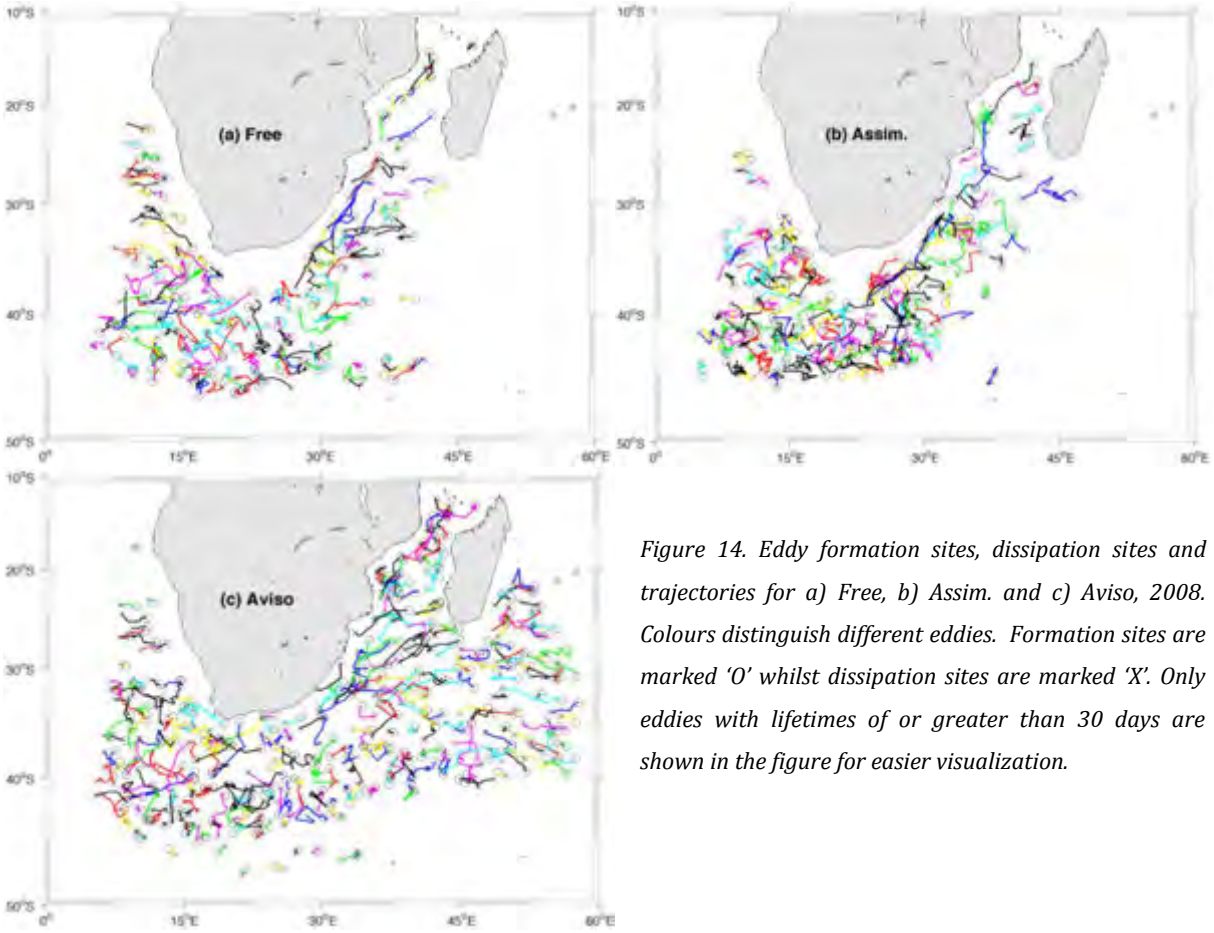

Figure 14. Eddy formation sites, dissipation sites and trajectories for a) Free, b) Assim. and c) Aviso, 2008. Colours distinguish different eddies. Formation sites are marked 'O' whilst dissipation sites are marked 'X'. Only eddies with lifetimes of or greater than 30 days are shown in the figure for easier visualization.

Analysis of the generation and dissipation sites, as well as eddy trajectories is a key component of this study. As a point of departure, birth and death sites and trajectories were mapped for eddies with lifetimes of or greater than 30 days as per Halo et al. (2014a; 2014b). A striking feature of the eddy density distribution analysis (Figure 13d) is the improvement in spatial distribution of eddies in the region south of Madagascar and the south-west Indian Ocean. However, this result seems at odds with the eddies and their trajectories shown in Figure 14. Though there are a few additional eddies in this region in Assim. compared to Free, they do not appear as numerous as might be expected given the greatly improved density distribution.

Figure 15 provides some insight into this issue, and prompted further analysis of not only the generation of eddies, but the coherence of their properties. By plotting the birth, death and propagation of *all* eddies (of any lifetime) for a given year, it becomes apparent that the model simulations initialize a number of eddies in the region south of Madagascar, but these eddies are not developed and propagated realistically. Furthermore, very few of these eddies have lifetimes greater than one week, which is why they fail to show up in the initial trajectory analysis (Figure 14). This is evident upon close inspection of Figure 15 which reveals that a number of eddies simulated by Free and Assim. have birth and dissipation sites which coincide spatially.

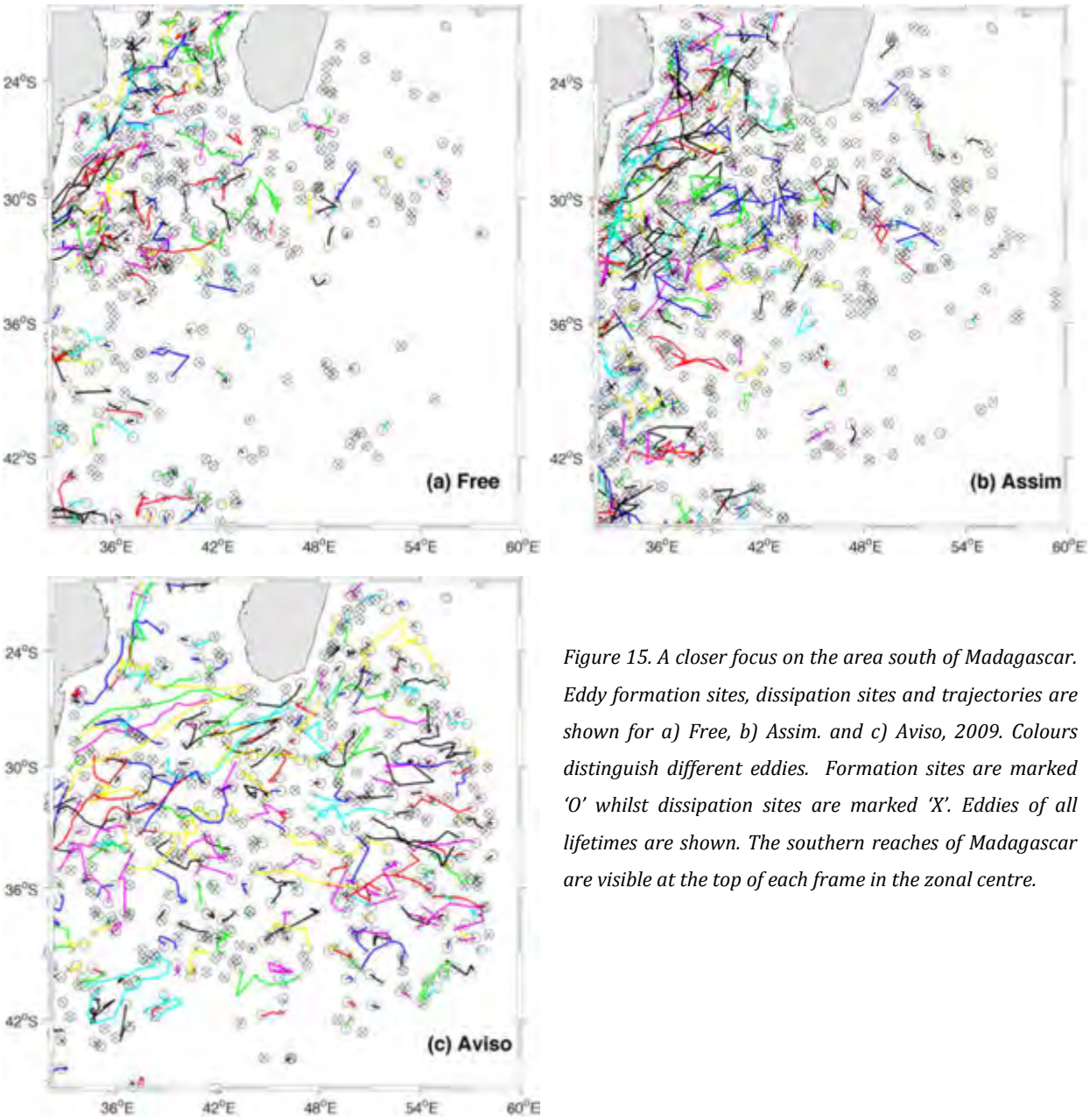

Figure 15. A closer focus on the area south of Madagascar. Eddy formation sites, dissipation sites and trajectories are shown for a) Free, b) Assim. and c) Aviso, 2009. Colours distinguish different eddies. Formation sites are marked 'O' whilst dissipation sites are marked 'X'. Eddies of all lifetimes are shown. The southern reaches of Madagascar are visible at the top of each frame in the zonal centre.

Figure 15 shows the relative improvement in eddy distribution of Assim. compared to Free, with Aviso as the reference. However, the development and realistic propagation of these features in the simulation appears concerning. One suggested reason for this phenomenon is that eddies with coincident birth/death sites have lifetimes of less than seven days (temporal resolution of the data), and are thus detected in only one timestep. As such, their propagation velocities cannot be inferred from their change in position between subsequent timesteps, and so they are shown to be generated and dissipated in their initial location. This was investigated, and found to account for the propagation anomalies seen in Figure 15. This is a limitation of the use of the tracking scheme with weekly data, and should be borne in mind for further analysis of results which rely on quantities inferred from propagation velocities (such as eddy kinetic energy and drift velocity). Upon investigation, these apparently "static" eddies were found to have otherwise coherent geometric and hydrographic properties, and can thus contribute to the analysis and comparison of eddy simulation.

Another possible explanation for this phenomenon is that the models are simulating inherently unrealistic eddy propagation in this region. They may imply unrealistically high translational velocities, and thus the detection and tracking scheme may classify the same eddy as two different eddies in subsequent timesteps. This would also result in the eddy in the initial timestep being assumed to have a lifetime < 7 days (no realistically corresponding eddy in following frame). Though this could be supposed to be a possible result of assimilation shock, where the input of external data into the numerical simulation might nudge the subsequent location of an eddy, implying a slightly excessive drift, it is unlikely to be the case. This is because the phenomenon occurs in both Free and Assim., and can thus not be conclusively linked to data assimilation.

As such, despite the anomalous propagation of these eddies having been shown to be an artefact of the tracking scheme combined with weekly-averaged data, it does point to an important aspect of the model solutions in this region. This is an unrealistic eddy lifetime, due to either (1) an unrealistically short-lived eddy signature in the SSH field, or (2) the unrealistic translational velocity of that signal, such that it is rejected by the tracking scheme. Moreover, it is unlikely that eddies in this region are so short-lived in reality, even if this is the reason for their insufficient tracking. The Aviso field shows a consistent propagation of long-lived eddies here. Furthermore, though sparse (especially-so to the south-east of Madagascar), the literature has shown at least some long-lived eddies to propagate fairly far in this region. Indeed, the tracks shown by Aviso in Figures 14 & 15 agree well with those in Figure 5 from de Ruijter et al.

(2004). The model would thus appear to require attention in this region if future configurations are to better resolve its mesoscale dynamics.

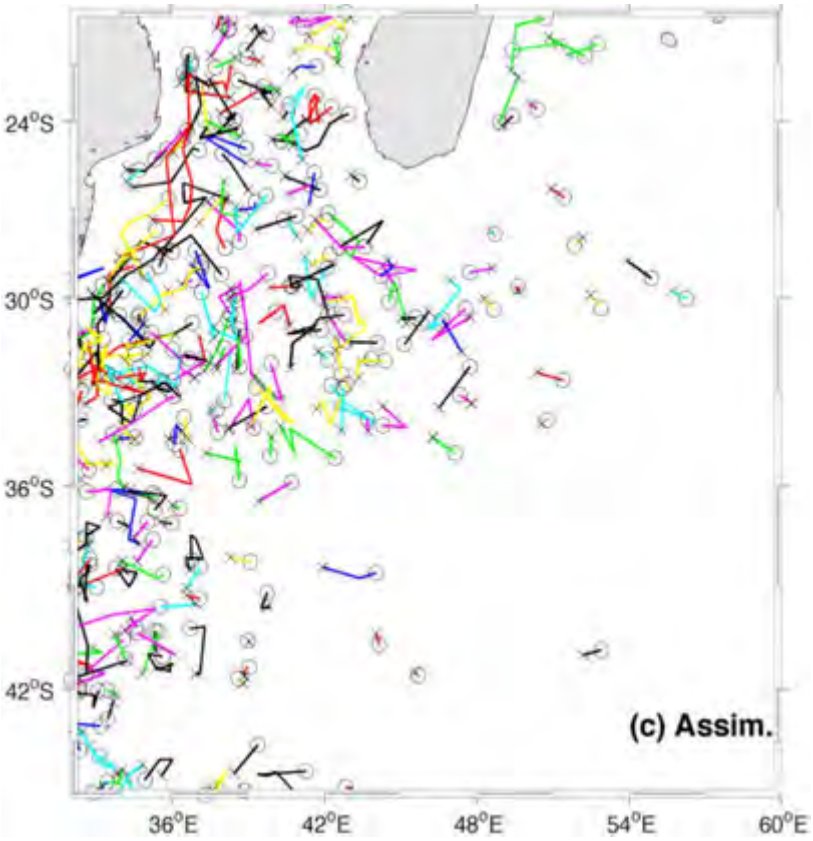

Figure 16. As shown in Figure 15 for Assim., but for eddies with minimum lifetimes of 7 days. With this minimum lifetime, eddies would be detected in at least two timestep frames, allowing for a realistic propagation velocity to be computed. Anomalous eddies (those showing no propagation) are absent.

Figure 16 shows the same results as before (Figure 15b), but excludes eddies with lifetimes < 7 days. Immediately apparent is the absence of the propagation anomalies which were previously evident. This supports the notion that these anomalies are eddies lasting less than a week, and can be addressed by filtering eddies with lifetimes < 7 days. However, even those eddy tracks which are visible in Figure 16 appear less realistic than those of Aviso, for the same region. Despite the lingering deficiency of eddies themselves, eddy tracks appear far shorter and less consistently directed than those in the corresponding Aviso field.

4.4 Eddy Geometry

4.4.1 Radius

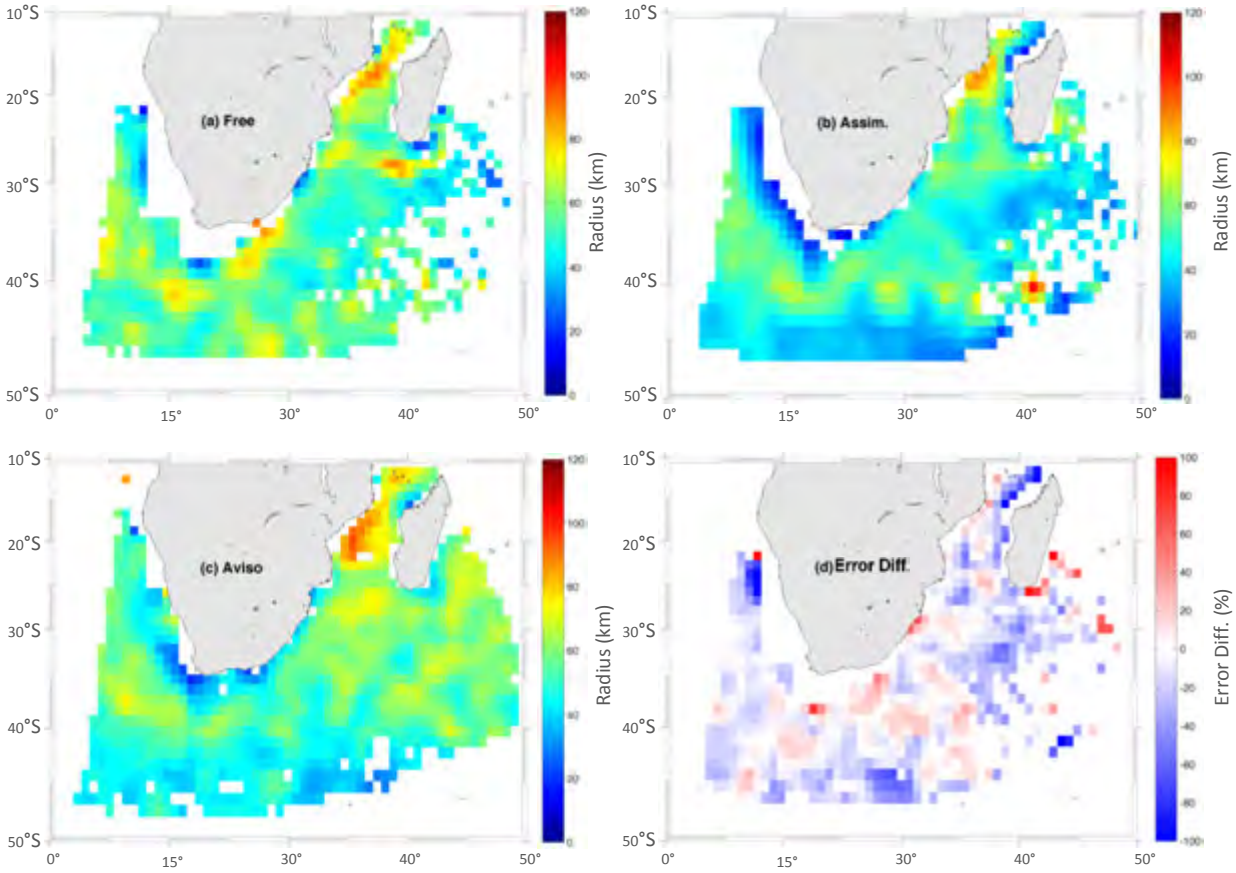

Figure 17. Eddy radius for a) Free, b) Assim. and c) Aviso. d) Shows relative error, with redder tones indicating closer adherence of Assim. to Aviso, and bluer areas indicating closer adherence of Free to Aviso.

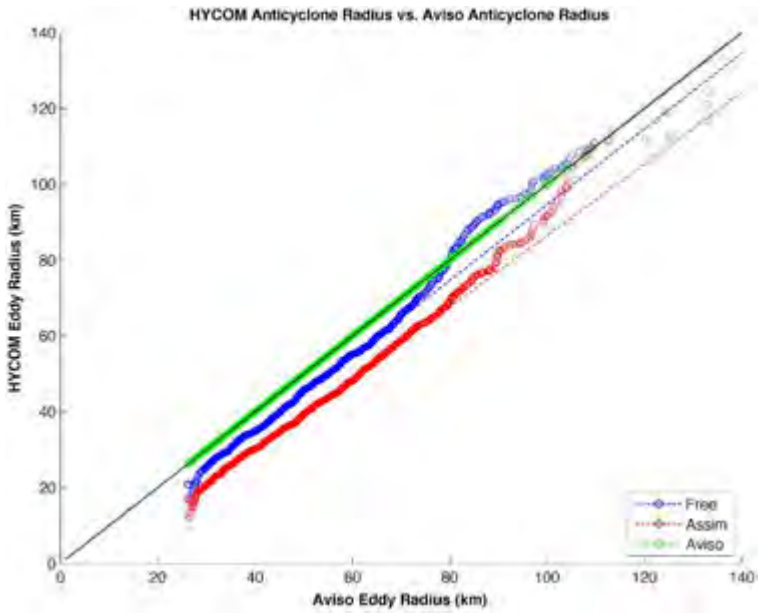

Figure 18. Quantile-quantile plot showing radii of Free (blue) and Assim. (red) anticyclones against radii of Aviso anticyclones. The solid black line indicates a 1:1 relation (i.e. identical to the reference, Aviso). Markers above (below) the solid black line indicate higher (lower) radii (than Aviso). Relationships between cyclone radii distributions are almost identical. Each marker represents the mean amplitude for that eddy. The red (blue) dotted line is the regression line for the Assim. (Free).

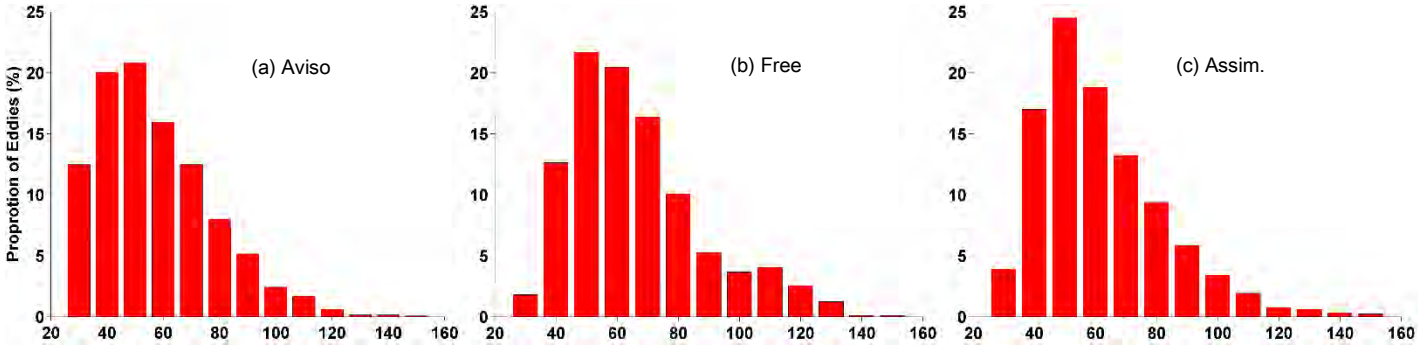

Figure 19. Anticyclone density distribution (%) as a function of radius for a) Aviso, b) Free and c) Assim.

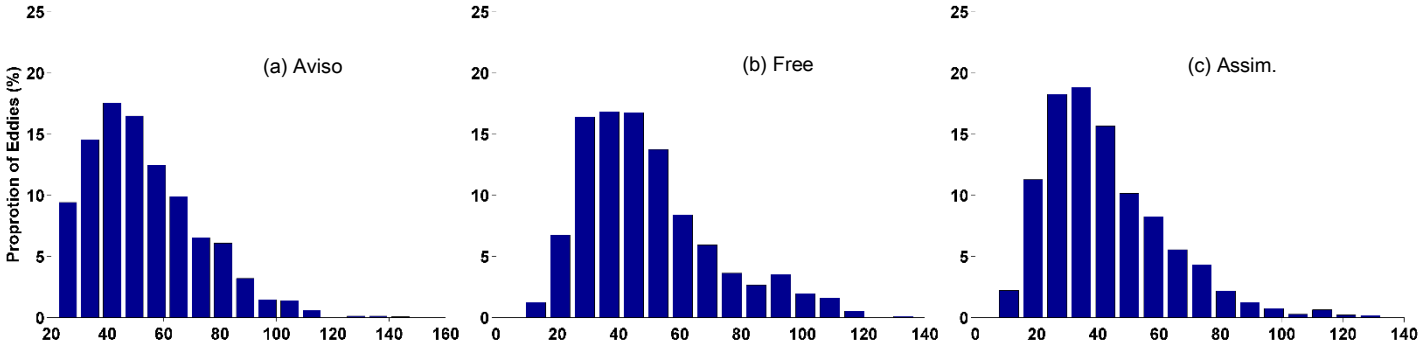

Figure 20. Cyclone density distribution (%) as a function of radius for a) Aviso, b) Free and c) Assim.

Figures 19 and 20 correspond to Figure 4, albeit for the entire Agulhas domain as opposed to just the Mozambique Channel. They indicate that anticyclones and cyclones display very similarly distributed radii in all three products. The distributions are all slightly positively skewed for both anticyclones and cyclones. Anticyclone and cyclone distributions peak at 25 km, 35 km and 35 km for Aviso, Free and Assim. respectively. This is somewhat different to the results of Halo et al. (2014a), which see Aviso and HYCOM anticyclones exhibiting bi-modally distributed radii, peaking at 30 km and 50 km in both. The corresponding cyclone distributions are normal, peaking at 70 km and 40 km respectively (Halo et al., 2014a). The difference in shape of anticyclone distributions may indicate a difference in the general geometric nature between anticyclones in the Mozambique Channel (as analysed in Halo et al., 2014a) and those in the rest of the Agulhas domain (those accounting for the distribution found in this study). The differences in peak lifetimes, however, are likely the result of the lifetimes of those eddies which Halo et al. (2014a) considered. Given the monotonous increase in eddy lifetime with radius, and the fact that only longer-lived (> 30 days) eddies were considered in that study, it follows that their radii would be distributed about higher peaks than the radii in this study.

4.4.2 Area

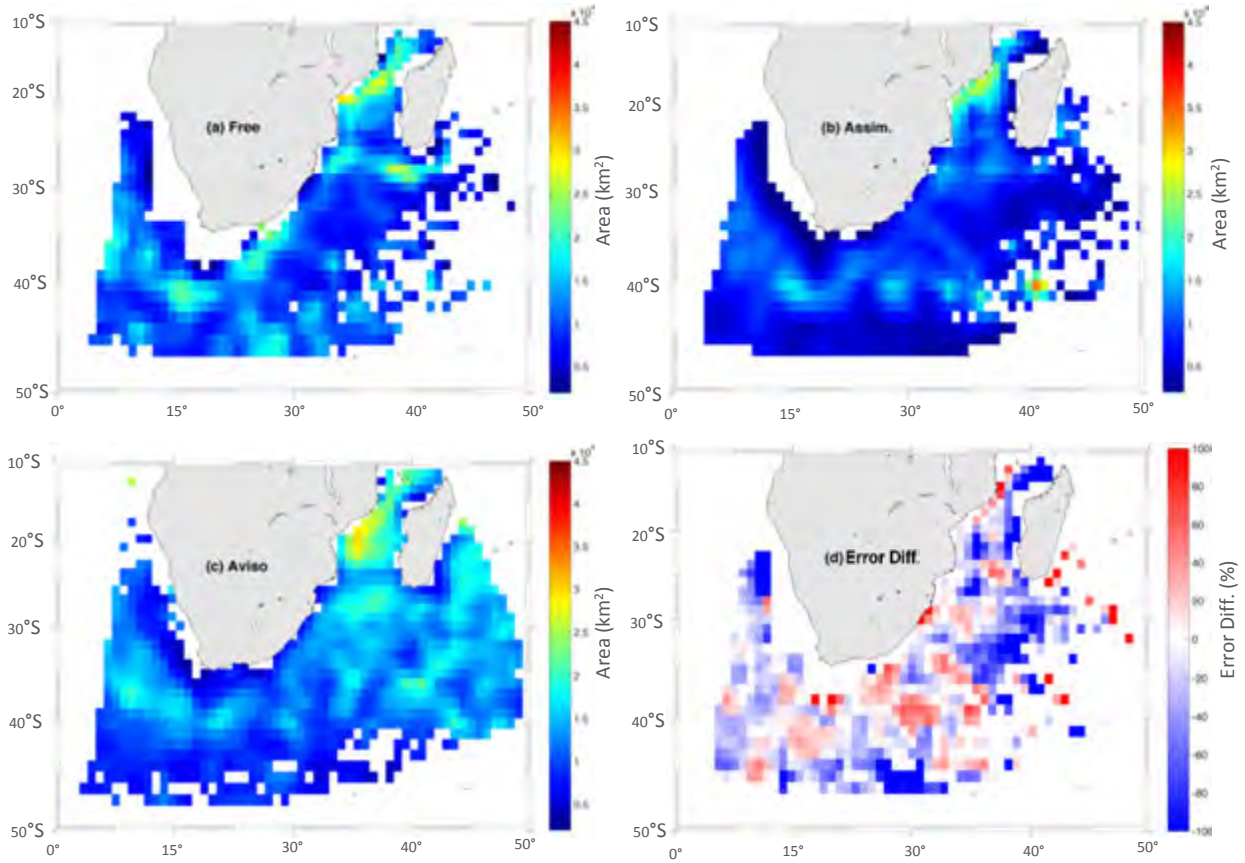

Figure 21. Eddy area for a) Free, b) Assim. and c) Aviso. d) Shows relative error, with redder tones indicating closer adherence of Assim. to Aviso, and bluer areas indicating closer adherence of Free to Aviso. As would be expected given the comparison of Radii in Figure 8., similar results are evident in eddy area.

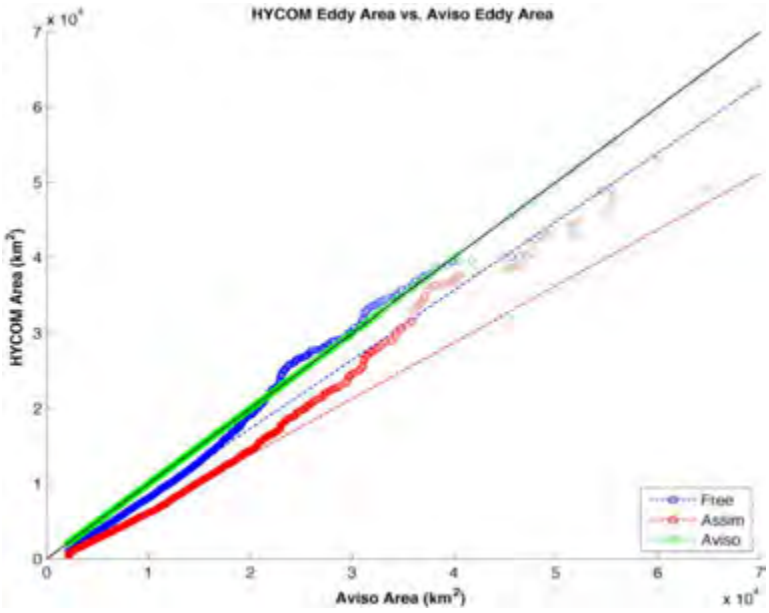

Figure 22. Quantile-quantile plot showing area of Free (blue) and Assim. (red) eddies against radii of Aviso eddies. The solid black line indicates a 1:1 relation (i.e. identical to the reference, Aviso). Markers above (below) the solid black line indicate higher (lower) radii (than Aviso). Each marker represents the mean area for that eddy. The red (blue) dotted line is the regression line for the Assim. (Free).

It appears that the overarching concern for the simulation of eddy geometry by Assim. is an underestimation of the size (radius) of eddies. All of the regions identified by the error maps as having higher error show smaller eddy radii. Figure 18 indicates that most of Free and Assim.'s radii distributions (highest density of blue and red markers) lie below that of Aviso (green reference line). However, radii of larger (> 80 km) eddies appear in better agreement in Free, but become overestimated at radii larger than this. Assim. radii remain underestimated throughout their distribution, with the exception of a handful of eddies with radii ~ 110 km.

This trend of underestimation carries through into the calculation of eddy area. However, at areas > 21, 000 km^2, Free shows a slight overestimation for the remainder of its distribution. Assim. areas remain underestimated throughout their distribution.

Degradation in accuracy or radius and area is noticeable in four main areas. These are to the east of Namibia, a region at the southern-most extent of the domain at around 30 ˚E, the general region south of Madagascar and the northerly reaches of the Mozambique Channel.

This result shares similarity to that of Halo et al. (2014a), which finds HYCOM to underestimate cyclone radii in the Mozambique Channel by 23 % when compared with other products. However, in this study, both anticyclone and cyclone radii are underestimated in both Free and Assim. throughout most of the domain. Furthermore, whilst there are isolated areas where Assim. appears more accurate than Free, Figure 18 indicates that these account for very few eddies out of the total distribution.

Interestingly, in Aviso, Free and Assim., both anticyclones and cyclones show right-skewed radii distributions. This is markedly different from the anticyclone and cyclone radii of Halo et al. (2014a), which are normally and bimodally distributed respectively. This reflects the intra-mesoscale range in size of eddies throughout the Agulhas domain (given that those in Halo et al. (2014a) are confined to the Mozambique Channel).

As such, the underestimation of eddy scale appears to be a recurring HYCOM artefact, which data assimilation has not addressed.

4.5 Eddy Dynamics

4.5.1 Amplitude

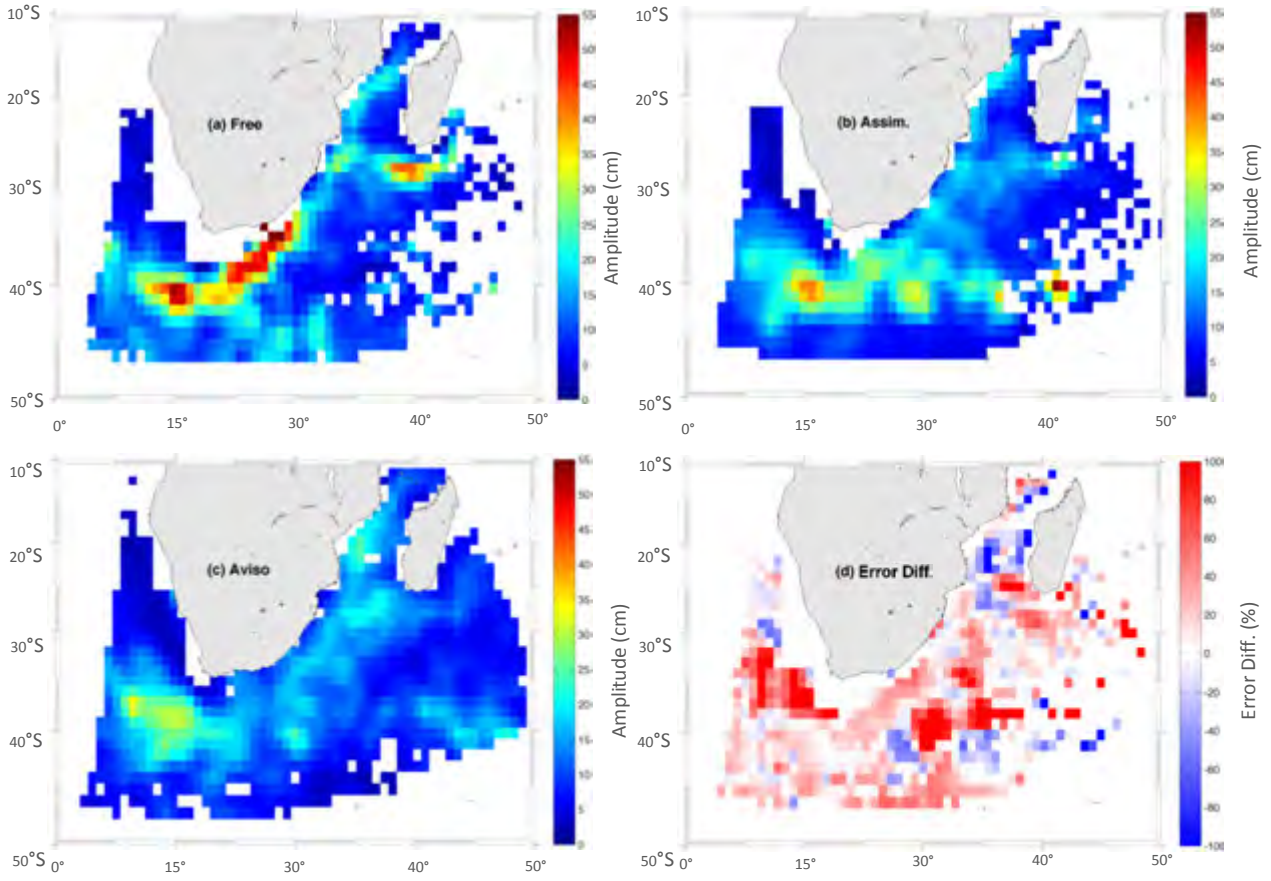

Figure 23. Eddy amplitude for a) Free, b) Assim. and c) Aviso. d) Shows relative error, with redder tones indicating closer adherence of Assim. to Aviso, and bluer areas indicating closer adherence of Free to Aviso.

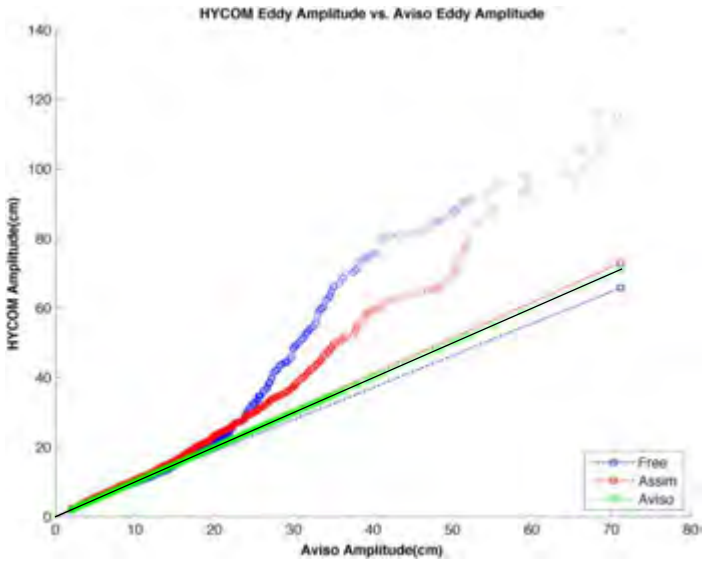

Figure 24. Quantile-quantile plot showing amplitudes of Free (blue) and Assim. (red) eddies against amplitudes of Aviso eddies. The solid black line indicates a 1:1 relation (i.e. identical to the reference, Aviso). Markers above (below) the solid black line indicate higher (lower) radii (than Aviso). Each marker represents the mean amplitude for that eddy. The red (blue) dotted line is the regression line for the Assim. (Free).

Continuous regions of moderate improvement, as well as three areas of high improvement are evident in the simulation of eddy dynamical properties by Assim. Eddy properties derived from Assim. in the shelf regions extending from south of Cape Agulhas in a north-westerly direction to west of Cape Columbine are in considerably better agreement with Aviso than those derived from Free. This is also true for the regions immediately south of Madagascar, as well as the open ocean domain extending from 36 S - 43 S and 28 E - 43 E (approximate centre of the study domain).

Conversely, immediately west of Madagascar appears to be an area where Assim. has yielded less accurate results.

For amplitudes up to 15 cm, Free and Assim. show similarly strong agreement with Aviso. However, for amplitudes greater than 22 cm, Assim. shows considerably better agreement than Free, though larger amplitudes in the distribution remain overestimated compared to Aviso. As a whole, Assim.'s distribution is more similar to that of Aviso.

4.5.2 Rotational Speed

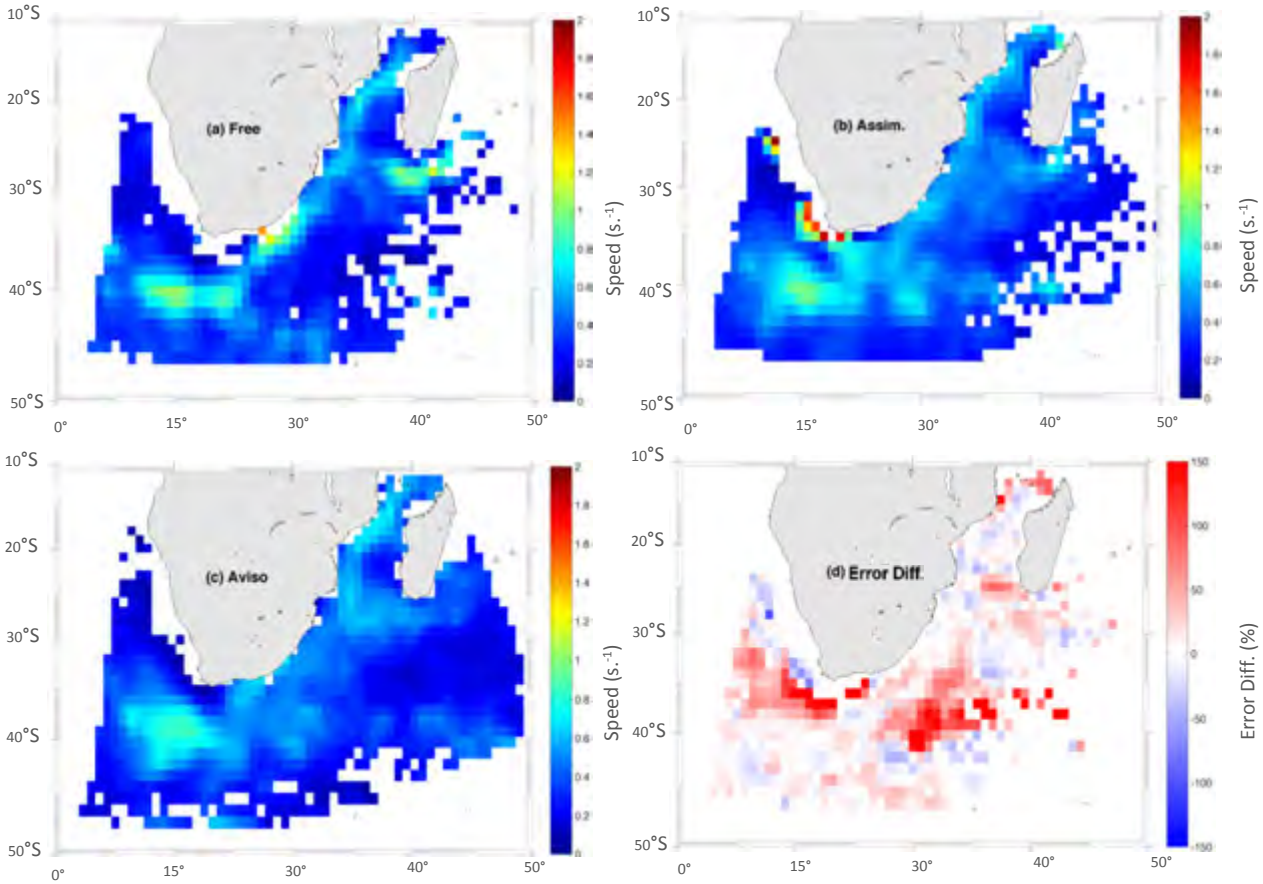

Figure 25. Maximum eddy rotational speed for a) Free, b) Assim. and c) Aviso. d) Shows relative error, with redder tones indicating closer adherence of Assim. to Aviso, and bluer areas indicating closer adherence of Free to Aviso.

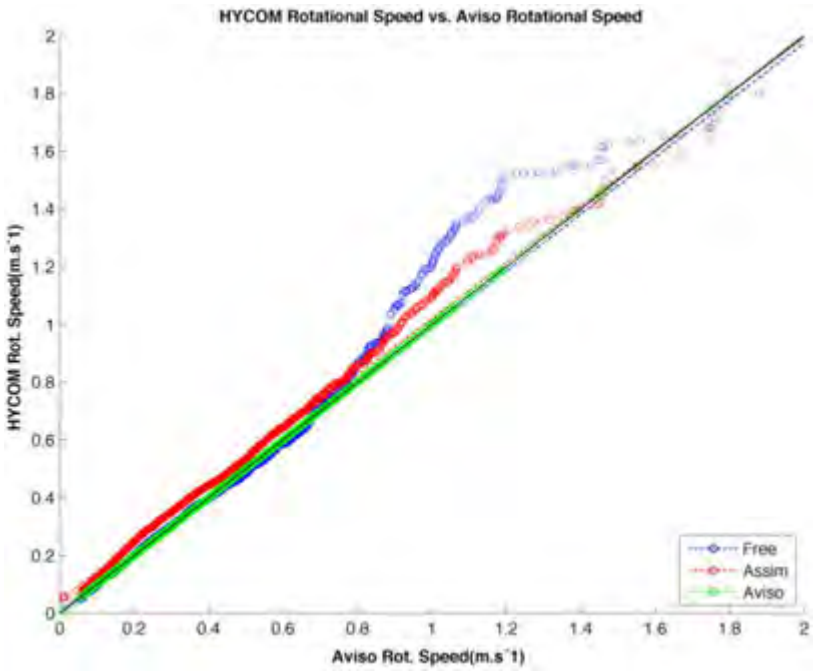

Figure 26. Quantile-quantile plot showing mean rotational speeds of Free and Assim. eddies against mean rotational speeds of Aviso eddies. The solid black line indicates a 1:1 relation (i.e. identical to the reference, Aviso). Markers above (below) the solid black line indicate higher (lower) rotational speeds (than Aviso). Each marker represents the mean rotational speed for that eddy. The red (blue) dotted line is the regression line for the Assim. (Free).

Eddy rotational speed is an important descriptive parameter for eddies, since it has a bearing on the advective capacity of eddies and lends insight into how energetic an eddy is. It can be thought of as the maximum geostrophic speed at which a particle trapped in an eddy might travel. The use of the maximum speed in this context is preferable for the subsequent computation of advective nonlinearity, as explained by Chelton et al. (2011). Assim. shows global improvement in rotational speed. Figure 25 indicates Assim.'s enhanced spatial similarity with respect to Aviso throughout the domain, with two regions of noticeably high improvement. These are in the same general locality as the areas of key improvement in the Amplitude analysis. Figure 26 shows that, though both simulations yield accurate mean rotational speeds across their distributions (evident in the regression lines, which indicate that the majority of their distributions are very similar), Assim. shows particular improvement in eddies with higher rotational speeds (> 0.9 $m.s^{-1}$).

Richardson (2007) finds the most energetic eddies, as computed from float and drifter data, to be those in the south-eastern regions of the Cape Basin (in this case, taken as extending only as far as far as 25 ˚E) (Richardson, 2007). This would appear to be in general agreement with the findings from this study, with the same region reflecting elevated rotational speeds relative to most of the domain. However, the highest rotational speeds, across all three products, are found to occur centred about 40 ˚S, 15 ˚E, and thus further west than in Richardson (2007). The link between decreasing rotational speed with distance from the retroflection region appears similarly valid in this study. Maximum speeds in this study appear higher than those derived from the drifter and float data, at 80-100 $cm.s^{-1}$, as opposed to the 40-60 $cm.s^{-1}$ of Richardson (2007). However, the small sample size (68) of the Richardson (2007) data should be borne in mind.

Naturally, as a consequence of the eddy distribution (as explained in Ch. 4.2 & 4.3), Free and Assim. are deficient in eddy statistics for the region south of Madagascar. Interestingly, however, those eddies which are detected in this region in Aviso exhibit relatively very low rotational speeds (~ 20 $cm.s^{-1}$). This indicates that, although there appear to be many eddies in the region, they are not very energetic.

4.5.3 Eddy Kinetic Energy

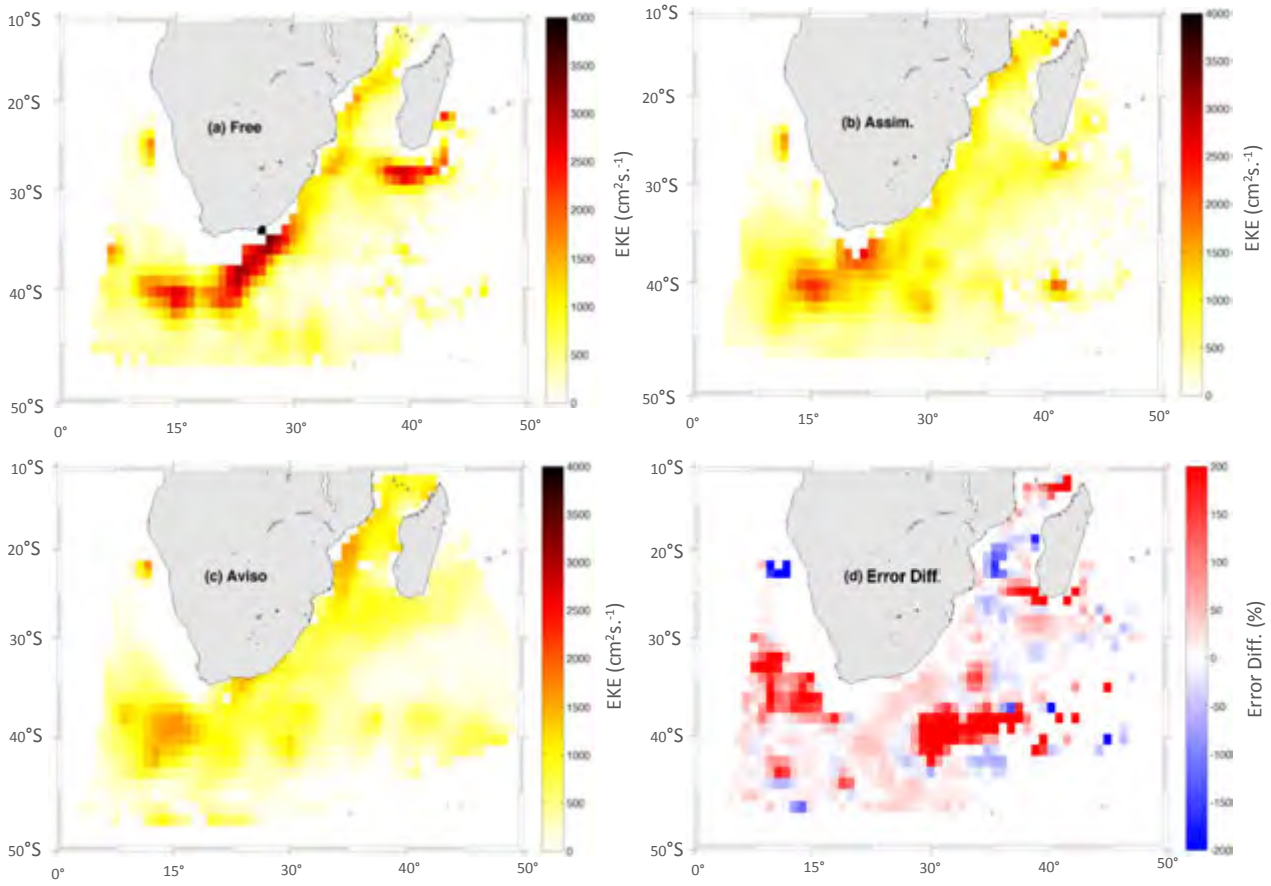

Figure 27. Eddy kinetic energy for a) Free, b) Assim. and c) Aviso. d) Shows relative error, with redder tones indicating closer adherence of Assim. to Aviso, and bluer areas indicating closer adherence of Free to Aviso.

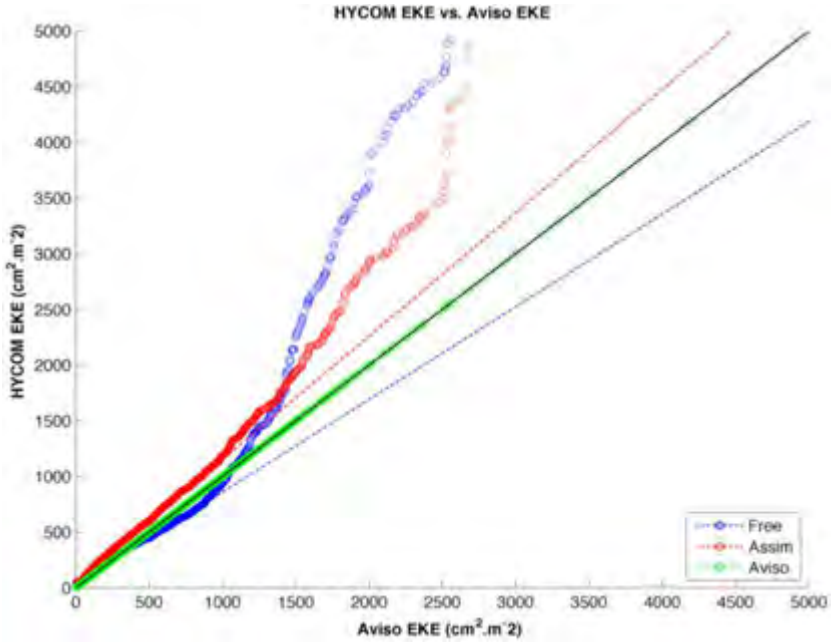

Figure 28. Quantile-quantile plot showing eddy kinetic energy of Free and Assim. eddies against rotational speeds of Aviso eddies. The solid black line indicates a 1:1 relation (i.e. identical to the reference, Aviso). Markers above (below) the solid black line indicate higher (lower) EKE (than Aviso). Each marker represents the mean EKE for that eddy. The red (blue) dotted line is the regression line for the Assim. (Free).

Analysis of eddy kinetic energy (EKE) is a useful component of ocean modelling and data assimilation validation (Backeberg et al., 2014). By providing insight into how chaotic the flow is, EKE can be used as a proxy in determining spatial mesoscale variability (Backeberg et al., 2014). Regions of higher (lower) EKE are subject to higher (lower) mesoscale variability. Assim. shows noticeable improvement in the spatial distribution of EKE over Free, in the same two areas as the dynamical improvements in Chapter 4.5.1, 4.5.2. Furthermore, excessive EKE along the southern extent of the Agulhas Current in Free is moderated in Assim., though it remains higher here, and southwest of Cape Agulhas than in Aviso. A noteworthy feature of the model results compared to Aviso is the lack of EKE in the region south of Madagascar and the southwest Indian Ocean. AGIOi, ARC112i, and INALT01i results in Loveday et al. (2014) show similarly low EKE compared to observation here, though it is difficult to tie this explicitly to a deficiency of eddies. In this study, Aviso shows at least some, albeit a low degree of EKE in this region. This is due to the deficiency of eddies in the model outputs. This result points to an important limitation of the HYCOM configurations for this region. Previous studies (e.g. Backeberg et al., 2014) have shown EKE from HYCOM in this region to be sufficient, and indeed at times, excessive. In such studies, EKE is calculated from continuous velocity fields. As such, whilst it might sometimes be the case, EKE cannot be explicitly linked to eddy activity, and could be the result of any chaotic flow or variation from the mean flow. In this study, computing EKE only where discrete eddies have been shown to exist (by virtue of their identification in the detection algorithm), allows us to attribute EKE in the region to mesoscale eddies. The results of this approach would suggest that, whilst the model appears to be simulating sufficient variability in the mean flow, it is not able to simulate and sustain eddies realistically. Interestingly even those eddies present in that region in Aviso are shown to have very low (~ 100 - 150 cm^2s^{-1}) EKE. The EKE in this region is the focus of Figure 29. This is a surprising result, given the number of Aviso eddies in this region and the coherence of their other properties evident in earlier results. Thus, though the models are not simulating a sufficiently high number of eddies in this region, the mesoscale variability shown by Aviso remains low despite the occurrence of eddies there.

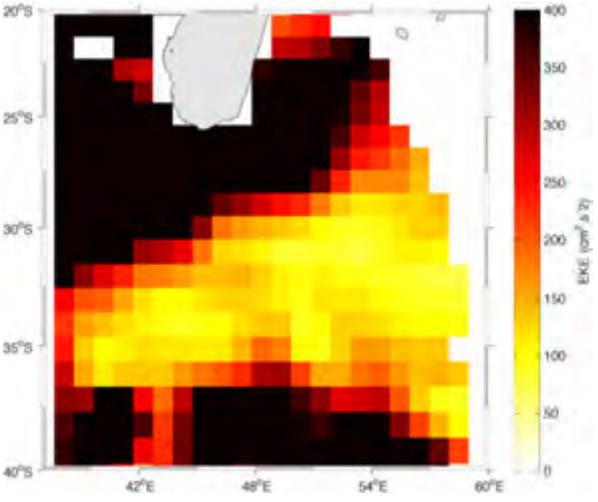

Figure 29. EKE for Aviso in the region south of Madagascar. Focussed scaling shows EKE in this region to be present, but low, indicating low mesoscale variability. In Free and Assim., this region is deficient in eddies, and thus their results do not reflect EKE in the region.

4.5.4 Advective Nonlinearity

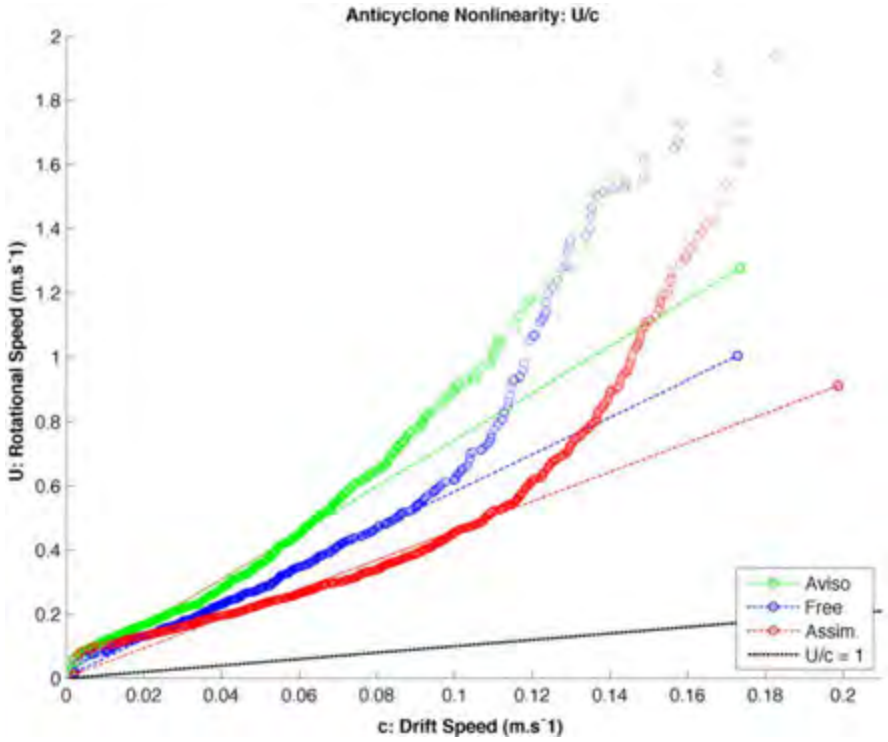

Figure 30. Quantile-quantile plot of advective nonlinearity (U/c) for Free, Assim. and Aviso anticyclones. Markers lying above the black dashed line (U/c = 1) indicate a dominance in rotational over translational speed of the flow. Coloured dashed lines are regression lines for their corresponding markers. The quantile plot of cyclone advective nonlinearity is almost identical, and not shown here. The red (blue) dotted line is the regression line for the Assim. (Free).

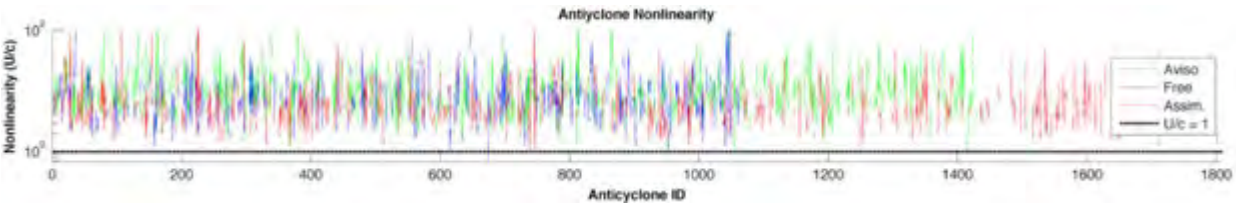

Figure 31. Anticyclone advective nonlinearity (U/c) against anticyclone identification number for Free, Assim. and Aviso. U/c = 1 is indicated by the dashed black line, and represents the threshold of nonlinearity above which an anticyclone is able to trap (and therefore advect) fluid. Nonlinearity (y-axis) is shown here on a logarithmic scale. The corresponding cyclone plot (not shown) is almost identical.

Advective nonlinearity is a useful parameter which allows the advective capacity of mesoscale features to be assessed. It is a non-dimensional ratio of rotational speed (U) against translational speed (c) (Chelton et al., 2011). Gaussian eddies exhibiting advective nonlinearities of $\frac{U}{c} > 1$ are able to trap fluid (Chelton et al., 2011) within their circulation and thus advect it according to their translational velocities away from that fluid's site of formation. This implies that such eddies are able to transport properties such as heat, salt and potential vorticity, as well as biogeochemical characteristics (Chelton et al., 2011). In the context of

Agulhas leakage and the advection of southeast Indian Ocean waters into the southeast Atlantic, this is an important property to consider.

Figure 30 shows the dominance of rotational speed over translational speed for nearly all anticyclones in all three data sets, implying that the overwhelming majority of anticyclones possess the important advective capacity. Aviso consistently shows the highest advective nonlinearity parameters for its anticyclones, followed by Free and then ASSM. This trend is mirrored for the nonlinearity parameters of Aviso, Free and Assim. cyclones, and is evident in the summary means provided in Table 4. Figure 31 confirms this trend by linking individual eddy identification numbers to their corresponding advective nonlinearity. Almost all data lie above the line $\frac{U}{c} = 1$, confirming that nearly all anticyclones are able to trap and advect fluids and their properties.

	Cyclones		**Anticyclones**	
	% Nonlinear	**Mean U/c**	**% Nonlinear**	**Mean U/c**
Aviso	99.8967	15.8508	99.8935	16.0174
Free	99.5448	10.4102	99.6805	10.4875
Assim.	99.8943	8.9196	98.4026	8.819

Table 4. Summary of advective nonlinearity for Aviso, Free and Assim. eddies. "Nonlinear" here is defined as having U/c > 1. An overwhelming majority of eddies (both cyclones and anticyclones) are nonlinear in all three datasets.

This result acts as a point of departure for an investigation into the reason behind Assim.'s consistently lower advective nonlinearity. An assessment of rotational and translational speeds independently from each other reveals that the ranking of nonlinearities can be attributed to a greater extent to consistent differences in translational speed than to differences in rotational speed. Assim. anticyclones exhibit slightly, but consistently higher translational speeds than those of Free and Aviso, whilst differences in rotational speeds are relatively minor in all cases.

4.5.5 Translational Velocity

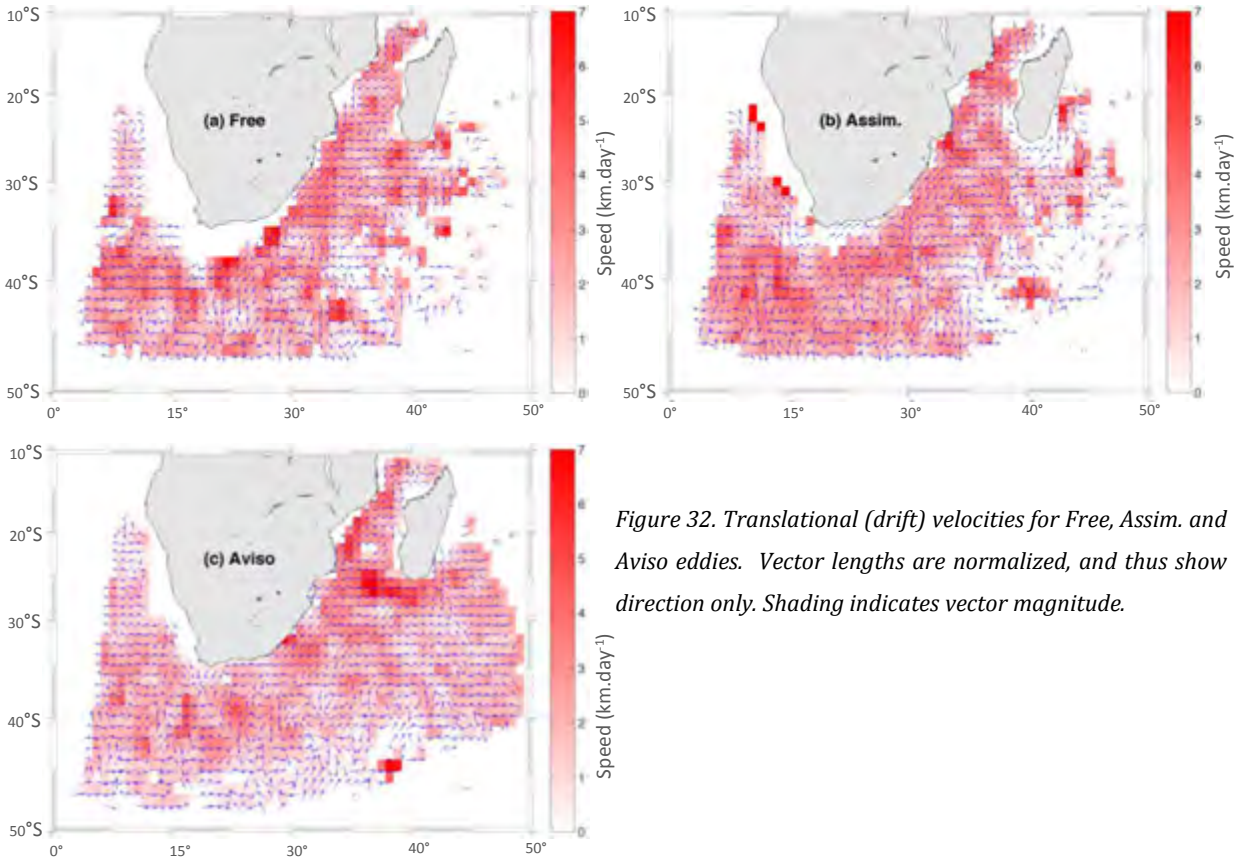

Figure 32. Translational (drift) velocities for Free, Assim. and Aviso eddies. Vector lengths are normalized, and thus show direction only. Shading indicates vector magnitude.

Figure 32 shows agreement with the expected (according to the literature) general propagation of mesoscale eddies for the study domain. This is the strong southward propagation of eddies in the Mozambique Channel (though this is too weak in Free), and a westward, turning north-westward propagation south and east of South Africa respectively. Southerly drift speeds appear improved in Assim. in the western part of the Mozambique Channel. Furthermore, excessive south-westerly drift speeds south and southeast of South Africa in Free appear moderated in Assim. A strong westward and south-westward propagation of eddies from south of Madagascar towards to east coast of south Africa is apparent in Aviso, but absent in the model outputs. However, Figure 32 shows the propagation of only a portion of the total eddies detected. This is because, an artefact of the tracking scheme, a number of eddies with otherwise coherent dynamical and geometric properties were assigned zero-translational velocities. This is case for 1040/2934 Aviso eddies (~ 35 %), 936/2190 Free eddies (~ 43 %) and 1876/3664 Assim. eddies (~ 51 %). This is most likely due to the temporal resolution of the data being insufficient for optimal tracking. In order for an eddy to be assigned a realistic translational velocity, it must necessarily appear as the same eddy in at least two consecutive frames. The distance travelled between the two frames is then divided by the time step between the two frames to give a translational speed (as per Ch. 3.3g). In this context, this implies, that only the

translational speeds of eddies with lifetimes ≥ 7 days can be computed, as the input SSH fields are weekly averages. As such, eddies with lifetimes < 7 days would be deemed to have no translation and instantaneous lifetimes. Cross-checking the number of eddies with zero-drift speeds and lifetimes < 7 days validates this as a probable cause. Of the 936 Free eddies with zero drift, 924 of them have instantaneous lifetimes (appearing in only one frame/timestep). The same relationship applies to 1828 Assim. eddies of the 1876 with zero drift speed, and 1046 Aviso eddies, with 1040 having zero drift speeds. Another cause may be excessive propagation or change in geometric/dynamical properties of eddies beyond the thresholds imposed by the tracking algorithm for the retention of an eddy's identity between timesteps. For example, if an eddy moves further than the tracking algorithm's accepted maximum displacement, it would be classified as a new eddy. Its predecessor would be deemed extinct, and assigned zero translational velocity if it was only present in that one timestep. This could explain the reason behind the relatively high prevalence of the phenomenon in Assim. (~ 51 %) compared to Free (~ 43 %) and Aviso (~ 35 %), in the context of assimilation shock. Assimilation shock occurs when the nudging of a solution closer to reality causes unrealistic or drastic changes in properties to occur. In nudging simulation results, data assimilation could change individual eddy properties between timesteps more than what is allowed by the tracking algorithm, causing discontinuity between timesteps. Confirmation of this causation/correlation would require further investigation and focussed analysis.

4.5.6 Summary

Eddy Property	RMSE		Relative RMSE (%)
	Free w.r.t. Aviso	Assim. w.r.t Aviso	
Cyclone Count	185.867	178.757	3.825
Anticyclone Count	187.852	190.074	-1.183
Area (km^2)	5890.400	6089.300	-3.377
Radius (km^2)	14.878	16.700	-12.251
Amplitude (cm)	10.351	6.577	36.458
Rot. Speed ($m.s^{-1}$)	0.215	0.156	27.480
Surface EKE ($m^2.s^{-2}$)	682.0649	588.9268	13.655
Anticyc. Vort. ($10e^{-5}.s^{-1}$)	0.6606	0.4976	24.675
Cyc. Vort. ($10e^{-5}.s^{-1}$)	0.4555	0.6207	-36.268

Table 5. Summary of computed RMSE of eddy properties for the two-year study period. Aviso results are taken to be the observational or reference value, with Free and Assim. as the predicted value in each case. Relative RMSE indicates which of the Free and Assim. configurations show closer agreement with Aviso. Where Assim. (Free) shows lower error with respect to Aviso, values are in green (red).

Chapter 5. Conclusions

This study successfully implements an objective method of quantifying the results of data assimilation by comparing free and assimilated configurations of HYCOM against each other, as well as against satellite altimetry data. It yields insight into mesoscale characteristics in the region, areas of concern and scope for possible future work in data assimilation and HYCOM modelling, and practical considerations for the implementation of an eddy detection and tracking scheme. Furthermore, certain limitations of the model are revealed only via the focussed analysis of eddy statistics as per this study (e.g. EKE), as opposed to standard output variables.

5.1 General Eddy Simulation & Density Distribution

Data assimilation yields a slight overall improvement in the stimulation of eddies, offering a closer total eddy count to that of Aviso than Free. Spatial analysis reveals that whilst Assim. offers better density distribution of cyclones, the spatial distribution of anticyclones is slightly degraded. Net spatial density distribution is improved.

The regions south of Madagascar, and the southwest Indian Ocean as far as 42 °S remain of concern in Free and Assim., showing a clear deficiency of eddies. Whilst Assim. offers improvement in this regard, it should be viewed in the context of the severe shortage of eddy simulation in this region in Free; thus the stimulation in Assim. of even a few eddies here yields large statistical improvement. The deficiency of eddies south of Madagascar and in the south-west Indian Ocean appears to be a spurious result, and may inherited from model boundary conditions. This requires further investigation.

5.2 Eddy Dynamics

Global improvement in eddy dynamics is evident in Assim. compared to Free. Four sites show consistent improvement across the range of computed eddy properties, with the exception of cyclone vorticity. These are two regions to the southwest and southeast of South Africa, one region in the northerly reaches of the Mozambique Channel, and the coastal region immediately south of Madagascar. Interestingly, this study highlights a deficiency of eddies (and therefore EKE) south of Madagascar. This contrasts with previous findings (eg. Backeberg et al., 2014) which suggest that HYCOM simulations overestimate EKE in this region. This difference may be due to the computation of EKE only where discrete eddies are present (as in this study) as

opposed to from continuous fields of geostrophic velocity, and highlights a possible caveat in the use of EKE as a proxy for mesoscale variability (i.e. representative of defined mesoscale eddy activity versus rather than arbitrary variation of the mean flow). This finding; that the limitation is only apparent when considering eddy statistics as opposed to standard model output variables; is important in directing future analysis of mesoscale dynamics and models' accuracy in simulating them.

It remains difficult to assess the impact of data assimilation on the translational velocity of eddies. This is most likely due to the combination of the temporal resolution of the data used in this study. Given that the majority of mesoscale eddies in the region are shorted lived (0-5 weeks), it may be of value to use daily data in future implementations of this detection and tracking scheme.

5.3 Eddy Geometry

Eddy radii and, consequently, area, are mostly underestimated in Free and Assim. when compared with Aviso. The underestimation is slightly more pronounced in Assim. At the larger end of eddy radii scales (> 80 km), Free begins to overestimate eddy radii. This trend carries through into the computation of eddy area. Free and Assim. eddies show consistently smaller areas until around 20, 000 km^2, after which Free overestimates eddy areas. Assim. continues to underestimate area throughout its distribution. As such, eddy geometry provides clear scope for improvement in future simulations.

Chapter 6. Limitations & Future Work

6.1 Temporal Resolution of Data

Weekly-averaged SSH data are used for input into the detection and tracking algorithm in this study. Whilst this is accepted as sufficient temporal resolution for the simulation of mesoscale eddies (Halo et al., 2014a; Backeberg et al., 2014), it may underrepresent eddies with lifetimes shorter than 7 days. These eddies, though short-lived, might nevertheless exhibit coherent and relevant dynamical properties; not least high rotational speeds, which would likely extend to non-linearity and therefore advective capacity. Such eddies are also relevant due to their prevalence throughout this study domain, as explained in Chapter 4.5, e. As such, future analyses would benefit from the use of increased temporal resolution. In this study, the impact is mitigated by the fact that weekly averages (rather than snapshots) are used, and so even short lived eddies might have some signature recorded in an averaged SSH field. Whether or not this signature would be sufficient to fulfil the criteria of the detection algorithm is questionable, however. The second issue arises in being unable to calculate propagation velocities for eddies which appear in only one time step. It would thus be beneficial to re-run this study using daily data.

6.2 Satellite Data as a Control Reference

The accuracy of Aviso altimetry has improved as a result of the incorporation of data from multiple sensors into composite products (Dohan & Maximenko, 2010). However, altimetry data are still subject to various limitations associated with temporal and spatial coverage, instrument precision and final gridded resolution. The merging process also incorporates at least some degree of uncertainty, associated with the mapping and interpolation of data from different sensors. Aviso conducts extensive calibration and processing to minimize the effects of these factors on the accuracy of the final MADT/MSLA products, and details this processing in the handbook available at http://www.aviso.altimetry.fr/fileadmin/documents/data/tools/hdbk_duacs.pdf. As such, whilst Aviso altimetry provides an excellent, practical reference against which to compare model data, they are not perfect as a control.

6.3 Individual Eddy Analysis

This study provides a general overview of longer-term eddy statistics, in order to determine the overall impact of data assimilation on the average mesoscale variability produced by different HYCOM configurations for a length of time. It does not compare individual, supposedly corresponding eddies with each other, between these configurations. The scope of such a comparison is beyond that of this project, but remains a worthwhile future objective in the evaluation of data assimilation in HYCOM. Though it remains unrealistic to expect different configurations to simulate exactly the same eddy (in time and space), a comparison of individual eddies from similar times and in the same location, and assessing full water column data (subsurface properties) might reveal further insight into the performance of the assimilated product.

6.4 Boundary Conditions South of Madagascar

A key result of this study is the deficiency of discrete eddies in the region south of Madagascar. This result provides a focal point for future HYCOM modelling of the Agulhas System. A study centred on this region, the boundary conditions it inherits, as well as the mesoscale activity would be worthwhile in attempting to enhance the accuracy of the modelling system.

References

Altimetry, A.V.I.S.O., 2009. SSALTO/DUACS user handbook: (M) SLA and (M) ADT near-real time and delayed time products. *Rep. CLS-DOS-NT*, *6*, p.51.

Backeberg, B.C., Bertino, L. and Johannessen, J.A., 2009. Evaluating two numerical advection schemes in HYCOM for eddy-resolving modelling of the Agulhas Current. *Ocean Science, 5,* pp.173-190.

Backeberg, B.C., Counillon, F., Johannessen, J.A. and Pujol, M.I., 2014. Assim.ilating along-track SLA data using the EnOI in an eddy resolving model of the Agulhas system. *Ocean Dynamics*, *64*(8), pp.1121-1136.

Backeberg, B.C., Johannessen, J.A., Bertino, L. and Reason, C.J., 2008. The greater Agulhas Current system: An integrated study of its mesoscale variability. *Journal of Operational Oceanography*, *1*(1), pp.29-44.

Backeberg, B.C., Penven, P. and Rouault, M., 2012. Impact of intensified Indian Ocean winds on mesoscale variability in the Agulhas system. *Nature Climate Change*, *2*(8), pp.608-612.

Beal, L.M., de Ruijter, W.P., Biastoch, A. and Zahn, R., 2011. On the role of the Agulhas system in ocean circulation and climate. *Nature*, *472*(7344), pp.429-436.

Barnier, B., Madec, G., Penduff, T., Molines, J.M., Treguier, A.M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J. and Derval, C., 2006. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean dynamics*, *56*(5-6), pp.543-567.

Biastoch, A., Böning, C.W. and Lutjeharms, J.R.E., 2008a. Agulhas leakage dynamics affects decadal variability in Atlantic overturning circulation. *Nature*, *456*(7221), pp.489-492.

Biastoch, A., Böning, C.W., Schwarzkopf, F.U. and Lutjeharms, J.R.E., 2009. Increase in Agulhas leakage due to poleward shift of Southern Hemisphere westerlies. *Nature*, *462*(7272), pp.495-498.

Biastoch, A., Lutjeharms, J.R.E., Böning, C.W. and Scheinert, M., 2008b. Mesoscale perturbations control inter-ocean exchange south of Africa. *Geophysical Research Letters*, *35*(20).

Bleck, R., 2002. An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean modelling*, *4*(1), pp.55-88.

Chassignet, E.P., Hurlburt, H.E., Smedstad, O.M., Halliwell, G.R., Hogan, P.J., Wallcraft, A.J., Baraille, R. and Bleck, R., 2007. The HYCOM (hybrid coordinate ocean model) data assimilative system. *Journal of Marine Systems*, *65*(1), pp.60-83.

Chelton, D.B., Deszoeke, R.A., Schlax, M.G., El Naggar, K. and Siwertz, N., 1998. Geographical variability of the first baroclinic Rossby radius of deformation. *Journal of Physical Oceanography*, *28*(3), pp.433-460.

Chelton, D.B., Schlax, M.G. and Samelson, R.M., 2011. Global observations of nonlinear mesoscale eddies. *Progress in Oceanography*, *91*(2), pp.167-216.

Counillon, F. and Bertino, L., 2009a. Ensemble Optimal Interpolation: multivariate properties in the Gulf of Mexico. *Tellus A*, *61*(2), pp.296-308.

Counillon, F. and Bertino, L., 2009b. High-resolution ensemble forecasting for the Gulf of Mexico eddies and fronts. *Ocean Dynamics*, *59*(1), pp.83-95.

Counillon, F. and Lisæter, K.A., NERSC-HYCOM 2.2.

de Ruijter, W.P., Brummer, G.J., Drijfhout, S.S., Lutjeharms, J.R.E., Peters, F., Ridderinkhof, H., Aken, H. and Leeuwen, P.J., 2006. Observations of the inter-ocean exchange around South Africa. *Eos, Transactions American Geophysical Union*, *87*(9), pp.97-101.

de Ruijter, W.P., Ridderinkhof, H., Lutjeharms, J.R., Schouten, M.W. and Veth, C., 2002. Observations of the flow in the Mozambique Channel.*Geophysical Research Letters*, *29*(10), pp.140-1.

de Ruijter, W.P., van Aken, H.M., Beier, E.J., Lutjeharms, J.R., Matano, R.P. and Schouten, M.W., 2004. Eddies and dipoles around South Madagascar: formation, pathways and large-scale impact. *Deep Sea Research Part I: Oceanographic Research Papers*, *51*(3), pp.383-400.

Dohan, K. and Maximenko, N., 2010. Monitoring ocean currents with satellite sensors. *Oceanography*, *23*(4), p.94.

Ducet, N., Le Traon, P.Y. and Reverdin, G., 2000. Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and-2. *Journal of Geophysical Research: Oceans*, *105*(C8), pp.19477-19498.

Evensen, G., 2003. The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), pp.343-367.

George, M.S., Bertino, L., Johannessen, O.M. and Samuelsen, A., 2010. Validation of a hybrid coordinate ocean model for the Indian Ocean. *Journal of Operational Oceanography*, *3*(2), pp.25-38.

Griffies, S. M., Böning, C., Bryan, F. O., Chassignet, E. P., Gerdes, R., Hasumi, H., and Webb, D., 2000. Developments in ocean climate modelling. *Ocean Modelling*, *2*(3), pp.123-192.

Halo, I., Backeberg, B., Penven, P., Ansorge, I., Reason, C. and Ullgren, J.E., 2014a. Eddy properties in the Mozambique Channel: A comparison between observations and two numerical ocean circulation models. *Deep Sea Research Part II: Topical Studies in Oceanography*, *100*, pp.38-53.

Halo, I., Penven, P., Backeberg, B., Ansorge, I., Shillington, F. and Roman, R., 2014b. Mesoscale eddy variability in the southern extension of the East Madagascar Current: Seasonal cycle, energy conversion terms, and eddy mean properties. *Journal of Geophysical Research: Oceans*, *119*(10), pp.7324-7356.

Henson, S.A. and Thomas, A.C., 2008. A census of oceanic anticyclonic eddies in the Gulf of Alaska. *Deep Sea Research Part I: Oceanographic Research Papers*, *55*(2), pp.163-176.

Isern-Fontanet, J., García-Ladona, E. and Font, J., 2006. Vortices of the Mediterranean Sea: An altimetric perspective. *Journal of Physical Oceanography*, *36*(1), pp.87-103.

Kalnay, E., 2003. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

Lathuiliere, C., Levy, M. and Echevin, V., 2010. Impact of eddy-driven vertical fluxes on phytoplankton abundance in the euphotic layer. *Journal of plankton research*, 0(0) pp.1-5.

Loveday, B.R., Durgadoo, J.V., Reason, C.J., Biastoch, A. and Penven, P., 2014. Decoupling of the Agulhas leakage from the Agulhas Current. *Journal of Physical Oceanography*, *44*(7), pp.1776-1797.

Lutjeharms, J. R. E., & Ansorge, I. J., 2001. The Agulhas return current. *Journal of Marine Systems*, *30*(1), pp. 115-138.

Lutjeharms, J.R.E. and Van Ballegooyen, R.C., 1988. The retroflection of the Agulhas Current. *Journal of Physical Oceanography*, *18*(11), pp.1570-1583.

Lutjeharms, J.R.E., De Ruijter, W.P.M. and Peterson, R.G., 1992. Interbasin exchange and the Agulhas retroflection; the development of some oceanographic concepts. *Deep Sea Research Part A. Oceanographic Research Papers*, *39*(10), pp.1791-1807.

Oke, P.R., Allen, J.S., Miller, R.N., Egbert, G.D. and Kosro, P.M., 2002. Assim.ilation of surface velocity data into a primitive equation coastal ocean model. *Journal of Geophysical Research: Oceans (1978–2012)*, *107*(C9), pp.5-1.

Oke, P.R., Sakov, P. and Corney, S.P., 2007. Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynamics*, *57*(1), pp.32-45.

Oke, P.R., Schiller, A., Griffin, D.A. and Brassington, G.B., 2005. Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Quarterly Journal of the Royal Meteorological Society*, *131*(613), pp.3301-3311.

Okubo, A., 1970, June. Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. *Deep Sea Research and Oceanographic Abstracts*, 17(3), pp. 445-454.

Omta, A.W., Llido, J., Garçon, V., Kooijman, S.A.L.M. and Dijkstra, H.A., 2009. The interpretation of satellite chlorophyll observations: The case of the Mozambique Channel. *Deep Sea Research Part I: Oceanographic Research Papers*, *56*(6), pp.974-988.

Pasquero, C., Provenzale, A. and Babiano, A., 2001. Parameterization of dispersion in two-dimensional turbulence. *Journal of Fluid Mechanics*, *439*, pp.279-303.

Peeters, F.J., Acheson, R., Brummer, G.J.A., De Ruijter, W.P., Schneider, R.R., Ganssen, G.M., Ufkes, E. and Kroon, D., 2004. Vigorous exchange between the Indian and Atlantic oceans at the end of the past five glacial periods. *Nature*, *430*(7000), pp.661-665.

Penven, P., Echevin, V., Pasapera, J., Colas, F. and Tam, J., 2005. Average circulation, seasonal cycle, and mesoscale dynamics of the Peru Current System: A modeling approach. *Journal of Geophysical Research: Oceans (1978–2012)*, *110*(C10).

Penven, P., Lutjeharms, J.R.E. and Florenchie, P., 2006. Madagascar: A pacemaker for the Agulhas Current system?. *Geophysical Research Letters*, *33*(17).

Reason, C.J.C., 2001. Evidence for the influence of the Agulhas Current on regional atmospheric circulation patterns. *Journal of Climate*, *14*(12), pp.2769-2778.

Richardson, P.L., 2007. Agulhas leakage into the Atlantic estimated with subsurface floats and surface drifters. *Deep Sea Research Part I: Oceanographic Research Papers*, *54*(8), pp.1361-1389.

Ridderinkhof, H. and De Ruijter, W.P.M., 2003. Moored current observations in the Mozambique Channel. *Deep Sea Research Part II: Topical Studies in Oceanography*, *50*(12), pp.1933-1955.

Ridderinkhof, H., Van der Werf, P.M., Ullgren, J.E., Van Aken, H.M., Van Leeuwen, P.J. and De Ruijter, W.P.M., 2010. Seasonal and interannual variability in the Mozambique Channel from moored current observations. *Journal of Geophysical Research: Oceans (1978–2012)*, *115*(C6).

Robinson, A.R. and Lermusiaux, P.F., 2000. Overview of data assimilation.*Harvard reports in physical/interdisciplinary ocean science*, *62*, pp.1-13.

Robinson, A.R., 1983. Overview and summary of eddy science. In *Eddies in marine science* (pp. 3-15). Springer Berlin Heidelberg.

Rouault, M., Penven, P. and Pohl, B., 2009. Warming in the Agulhas Current system since the 1980's. *Geophysical Research Letters*, *36*(12).

Schouten, M.W., de Ruijter, W.P. and van Leeuwen, P.J., 2002. Upstream control of Agulhas Ring shedding. *Journal of Geophysical Research: Oceans (1978–2012)*, *107*(C8), pp.23-1.

Souza, J.M.A.C.D., De Boyer Montegut, C. and Le Traon, P.Y., 2011. Comparison between three implementations of automatic identification algorithms for the quantification and characterization of mesoscale eddies in the South Atlantic Ocean. *Ocean Science*, *7*(3), pp.317-334.

Srinivasan, A., Chassignet, E.P., Bertino, L., Brankart, J.M., Brasseur, P., Chin, T.M., Counillon, F., Cummings, J.A., Mariano, A.J., Smedstad, O.M. and Thacker, W.C., 2011. A comparison of sequential assimilation schemes for ocean prediction with the Hybrid Coordinate Ocean Model (HYCOM): Twin experiments with static forecast error covariances. *Ocean Modelling*,*37*(3), pp.85-111.

Stramma, L. and Lutjeharms, J.R., 1997. The flow field of the subtropical gyre of the South Indian Ocean. *Journal of Geophysical Research: Oceans (1978–2012)*, *102*(C3), pp.5513-5530.

Swart, N.C., Lutjeharms, J.R.E., Ridderinkhof, H. and De Ruijter, W.P.M., 2010. Observed characteristics of Mozambique Channel eddies. *Journal of Geophysical Research: Oceans (1978–2012)*, *115*(C9).

Talagrand, O., 1997. Assimilation of observations, an introduction. *Journal-Meteorological Society of Japan Series 2*, *75*, pp.81-99.

Weimerskirch, H., Le Corre, M., Jaquemet, S., Potier, M. and Marsac, F., 2004. Foraging strategy of a top predator in tropical waters: great frigatebirds in the Mozambique Channel. *Marine Ecology Progress Series*, *275*, pp.297-308.

Weiss, J., 1991. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica D: Nonlinear Phenomena*, *48*(2), pp.273-294.

Xie, J., Counillon, F., Zhu, J. and Bertino, L., 2011. An eddy resolving tidal-driven model of the South China Sea assimilating along-track SLA data using the EnOI. *Ocean Science*, *7*(5), pp.609-627.

***